Epidemics and networks

Alain Barrat

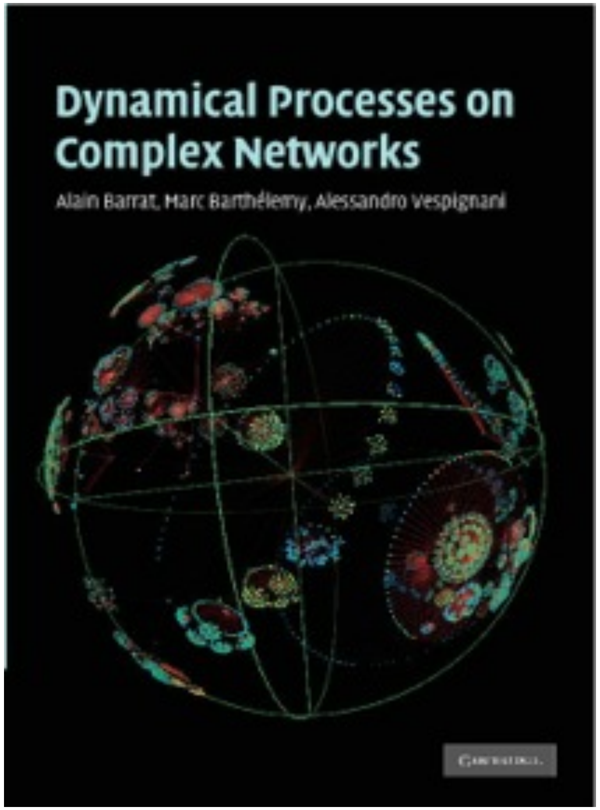
CPT, Marseille, France

ISI, Turin, Italy

cnrs
dépasser les frontières

CPT
Centre de Physique Théorique

INSTITUTE
FOR SCIENTIFIC INTERCHANGE
FOUNDATION


Dynamical Processes on Complex Networks
Alain Barrat, Marc Barthélemy, Alessandro Vespignani

http://www.cpt.univ-mrs.fr/~barrat
http://www.cxnets.org
http://www.sociopatterns.org

ECCS 2014 26 Jun
@eccs14

Remember to register asap and to book a place. Please tell us the satellites you'll follow. We need to arrange the conference rooms :)

Outline

I. **Models**
II. **Mean-field picture**
III. **Degree-based mean-field**
IV. **Individual-based mean-field**
V. **Numerics**
VI. **Epidemics in meta-population networks**
VII. **Epidemics in multiplex networks**
VIII. **Epidemics in temporal networks**

Epidemiology

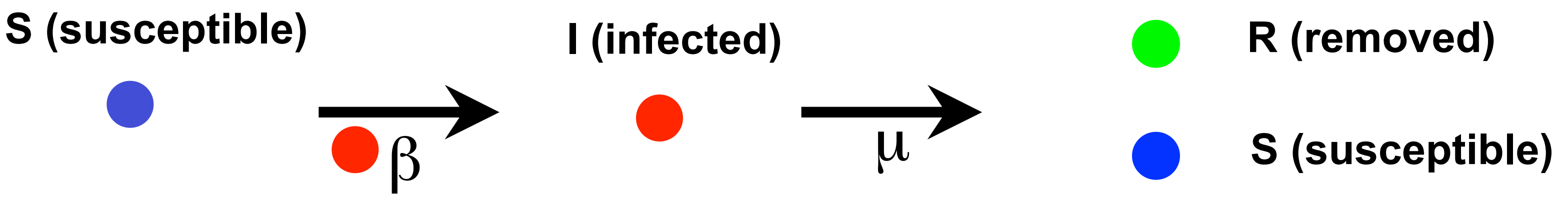
Two levels:

- Microscopic: researchers try to disassemble and kill new viruses => quest for vaccines and medicines

- Macroscopic: statistical analysis and modeling of epidemiological data in order to find information and policies aimed at lowering epidemic outbreaks => macroscopic prophylaxis, vaccination campaigns…

Standard epidemic modeling

<u>Compartments:</u> S, I, R...



S (susceptible)

β

I (infected)

μ

R (removed)

S (susceptible)

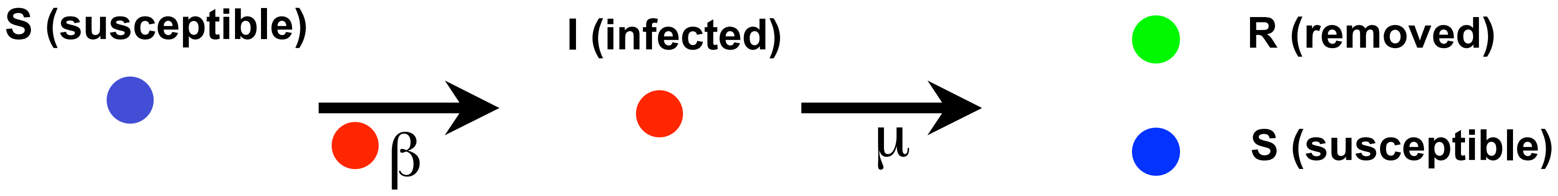
Neglecting differences in:

- age
- gender
- health
- social class/status
- …

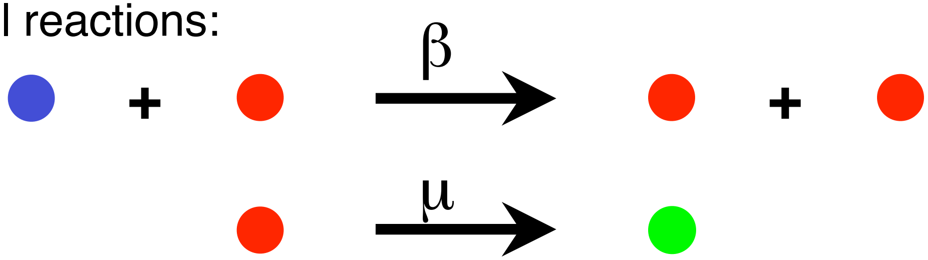
- susceptibility to disease
- latency
- severity of disease
- …

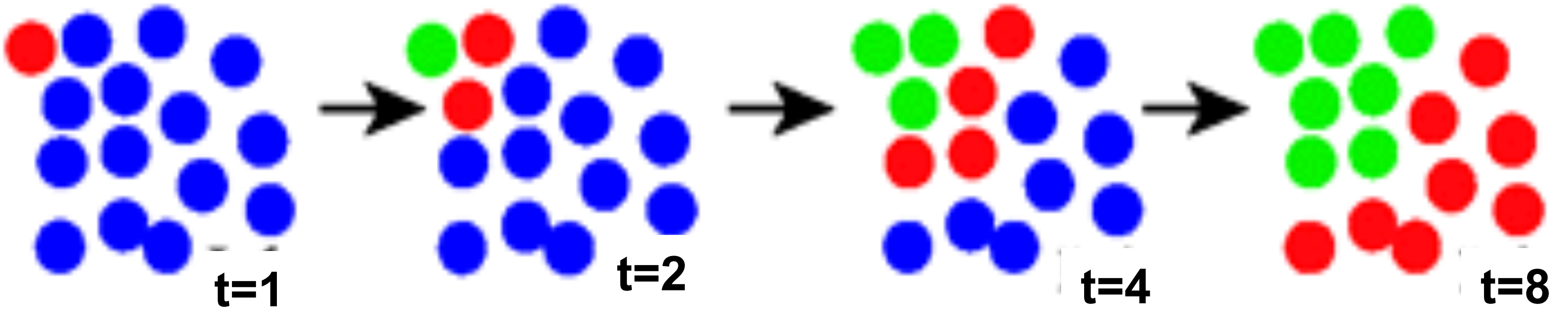
Standard epidemic modeling

Compartments: **S, I, R...**



Similar to chemical reactions:





Standard epidemic modeling

SI

$$S + I \xrightarrow{\beta} I + I$$

SIS

$$S + I \xrightarrow{\beta} I + I$$

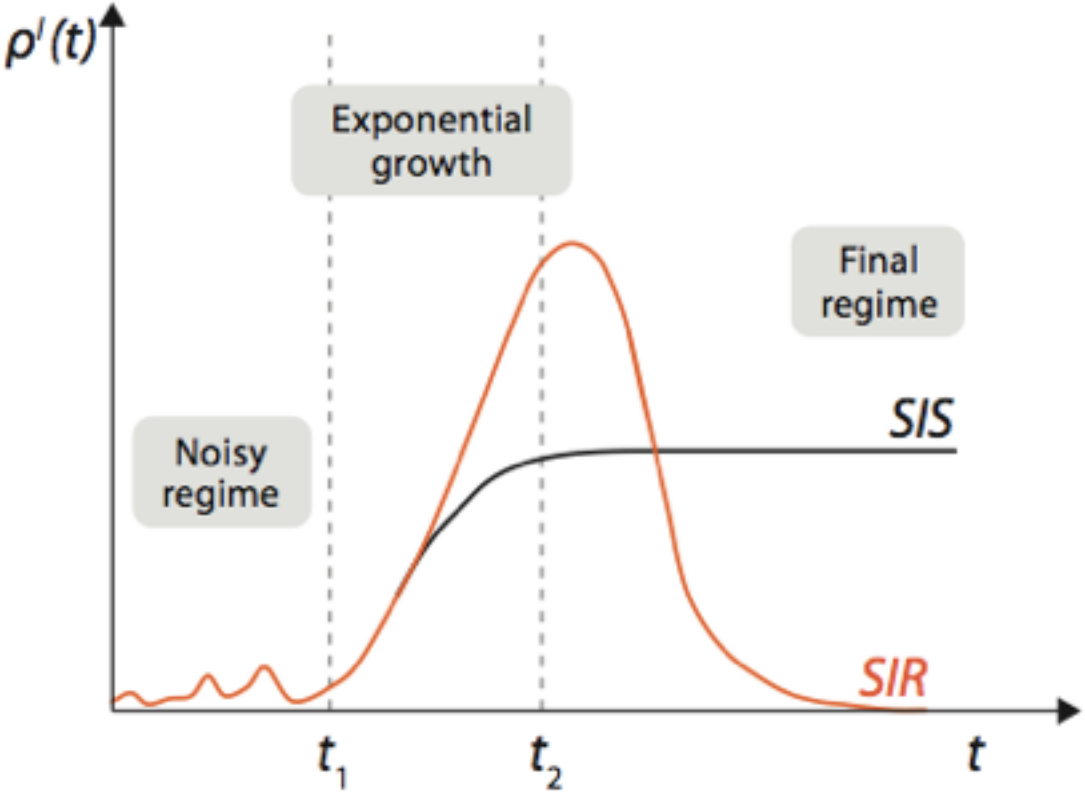
$$I \xrightarrow{\mu} S$$

SIR

$$S + I \xrightarrow{\beta} I + I$$

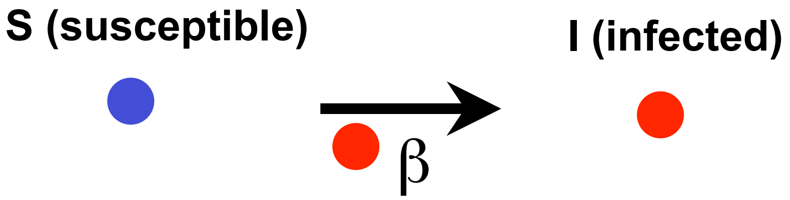
$$I \xrightarrow{\mu} R$$

Stages of an epidemic outbreak: population level



Infected individuals => prevalence/incidence

Transmission



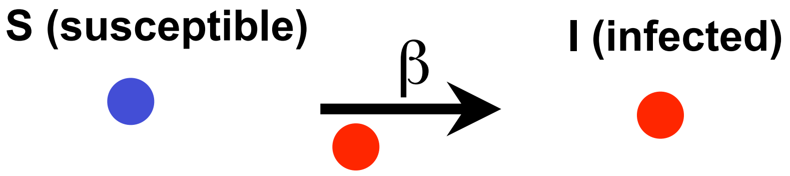
HOMOGENEOUS MIXING ASSUMPTION

Individual in state S, with k contacts, among which n infectious: in the homogeneous mixing approximation, the probability to get the infection in each time interval dt is:

$$
\begin{aligned}
\text{Proba}(S \rightarrow I) &= 1 - \text{Proba(not to get infected by any infectious)} \\
&= 1 - (1 - \beta\, dt)^{n} \\
&\cong \beta\, n\, dt \quad (\beta\, dt << 1) \\
&\cong \beta\, k\, i\, dt \quad \text{as} \quad n \sim k\, i \textit{ for homogeneous mixing}
\end{aligned}
$$

Hypothesis of mean-field nature:
every individual sees the same density of infectious among his/her contacts, equal to the average density in the population

The SI model


S (susceptible)

β

I (infected)

N individuals

$I(t)$=number of infectious, $S(t)=N-I(t)$ number of susceptible

$i(t)=I(t)/N$, $s(t)=S(t)/N = 1- i(t)$

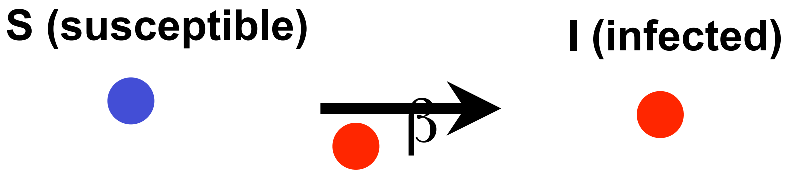
If k = $<k>$ is the same for all individuals (homogeneous network):

$$\frac{dI}{dt} = S(t) \times Proba(S \rightarrow I)$$

$$= \beta k S(t) i(t)$$

$$\boxed{\frac{di}{dt} = \beta k i(t)(1 - i(t))}$$

The SI model



N individuals

I(t)=number of infectious, *S(t)*=*N*-*I(t)* number of susceptible

i(t)=*I(t)*/*N* , *s(t)*=*S(t)*/*N*

$$\frac{di}{dt} = \beta\langle k\rangle i(1-i)$$

$$i(t) = \frac{i_0\exp(t/\tau)}{1+i_0(\exp(t/\tau)-1)} \qquad \tau = 1/(\beta\langle k\rangle)$$

The SIS model

S (susceptible)

I (infected)

S (susceptible)

β

μ

N individuals

I(t)=number of infectious, *S(t)=N-I(t)* number of susceptible

i(t)=I(t)/N , s(t)=S(t)/N

Homogeneous mixing

$$\frac{di}{dt} = \beta \langle k \rangle i(1-i) - \mu i$$

Competition of two time scales: 1/μ and 1/(β <k>)

The SIR model

N individuals

I(t)=number of infectious, *S(t)* number of susceptible, *R(t)* recovered

$i(t)=I(t)/N$, $s(t)=S(t)/N$, $r(t)=R(t)/N=1-i(t)-s(t)$

Homogeneous mixing:

$$\frac{ds}{dt} = -\beta\langle k\rangle i(t)s(t)$$

$$\frac{di}{dt} = \beta\langle k\rangle i(t)s(t) - \mu i(t)$$

$$\frac{dr}{dt} = \mu i(t)$$

Competition of two time scales: $1/\mu$ and $1/(\beta <k>)$

SIS and SIR models: linear approximation

Short times, $i(t) \ll 1$ (and $r(t) \ll 1$ for the SIR)

$$\frac{di}{dt} \approx (\beta\langle k\rangle - \mu)i(t)$$

Exponential evolution $\exp(t/\tau)$, with

$$\boxed{1/\tau = \beta\langle k\rangle - \mu}$$

If $\beta\langle k\rangle > \mu$: exponential growth

If $\beta\langle k\rangle < \mu$: extinction

Epidemic threshold condition: $\beta\langle k\rangle = \mu$

Long time limit, SIS model

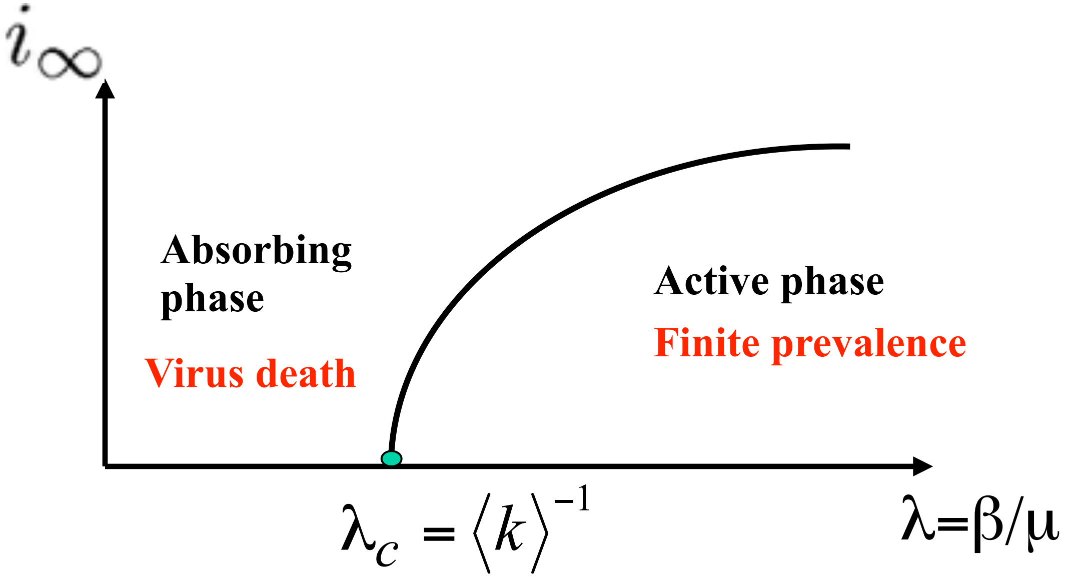
Stationary state: *di/dt = 0* $\mu i_\infty = \beta \langle k \rangle i_\infty (1 - i_\infty)$

$$\beta\langle k\rangle < \mu \Rightarrow i_\infty = 0$$

Epidemic threshold condition: $\beta\langle k\rangle = \mu$

$$\beta\langle k\rangle > \mu \Rightarrow i_\infty = 1 - \mu/(\beta\langle k\rangle)$$

Phase diagram:



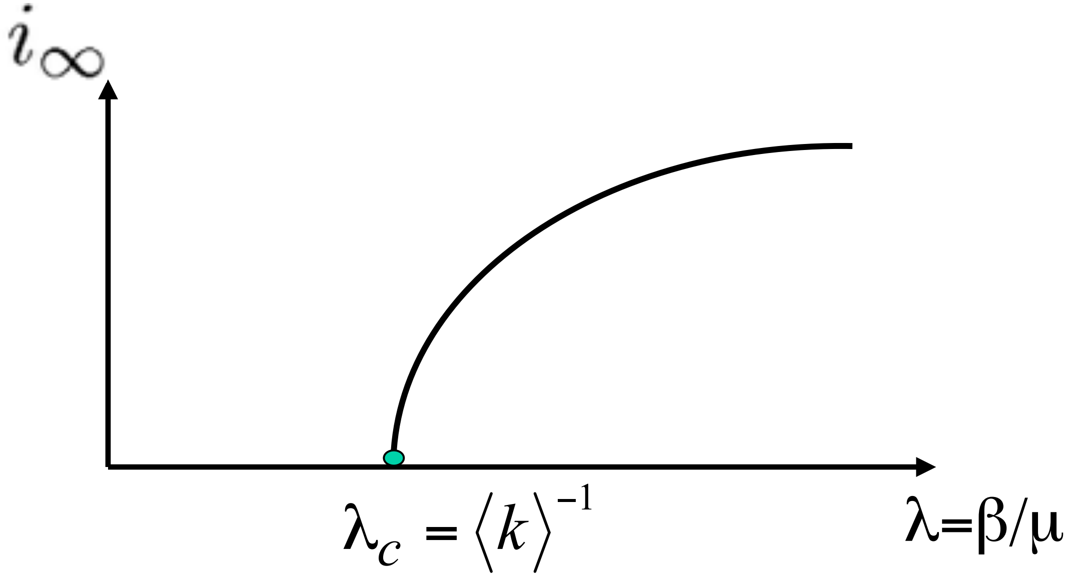
Immunization

Fraction *g* of immunized (vaccinated) individuals:

reduce population of susceptible individuals

$$S \to S \times (1-g)$$

$$\frac{di}{dt} = \beta k i(t)(1-i(t))(1-g)$$



Immunization

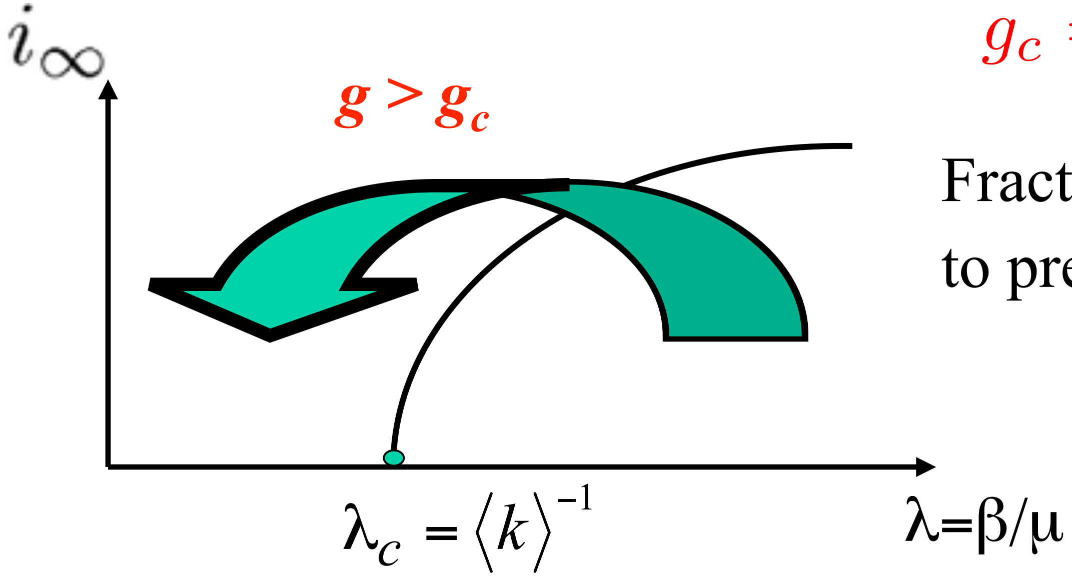
Fraction *g* of immunized (vaccinated) individuals:
reduce population of susceptible individuals

Equivalent to a reduction of β:

$$\beta \to (1-g)\beta$$
$$\lambda \to (1-g)\lambda$$

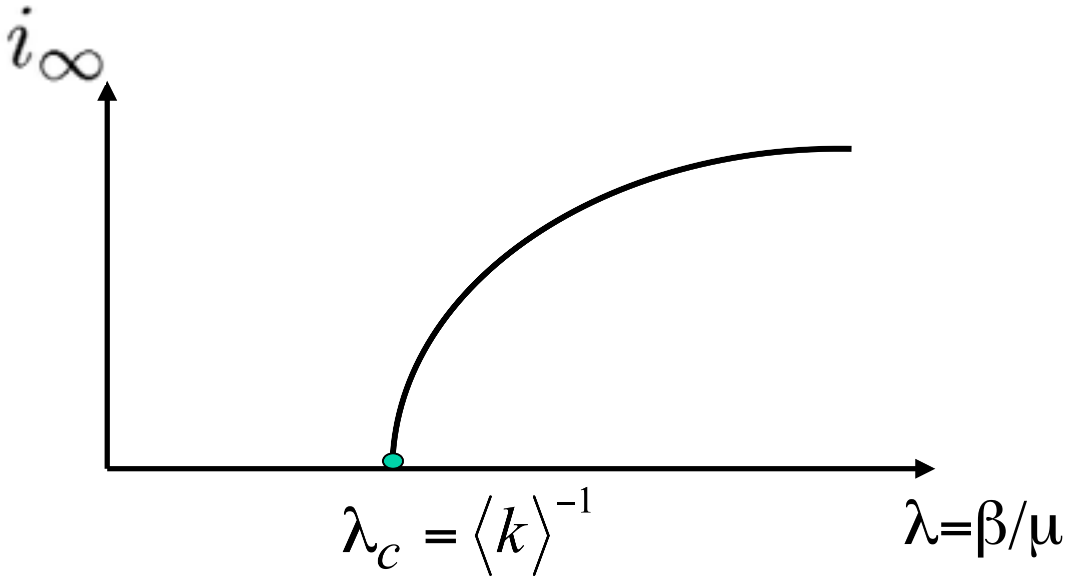
=> critical immunization threshold

$$g_c = 1 - \mu/(\beta\langle k\rangle)$$



Fraction of population to vaccinate
to prevent an outbreak

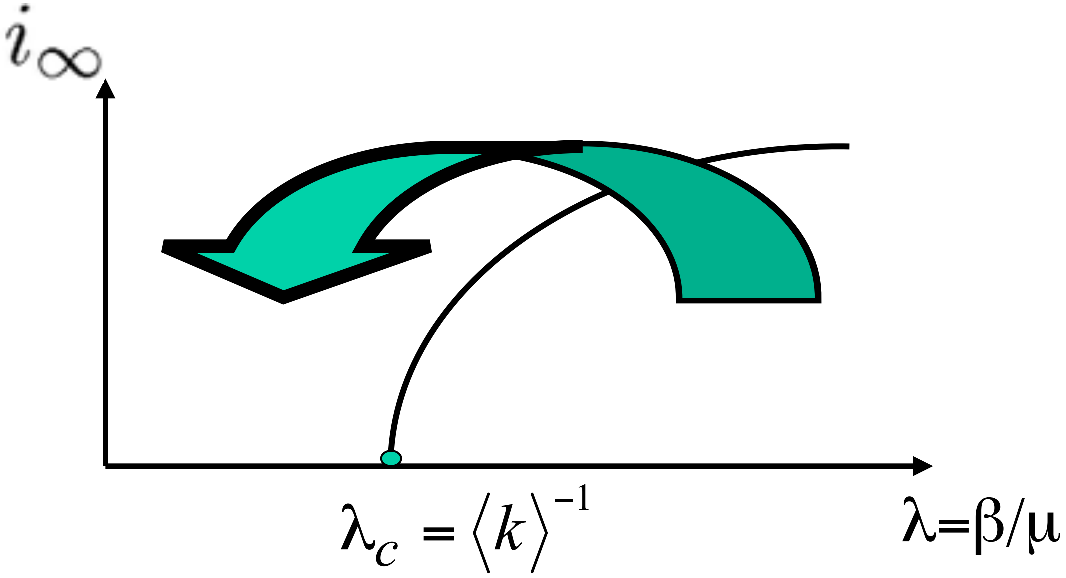
Homogeneous mixing: summary



Competition of time scales

=> Epidemic threshold condition: $\beta\langle k\rangle = \mu$

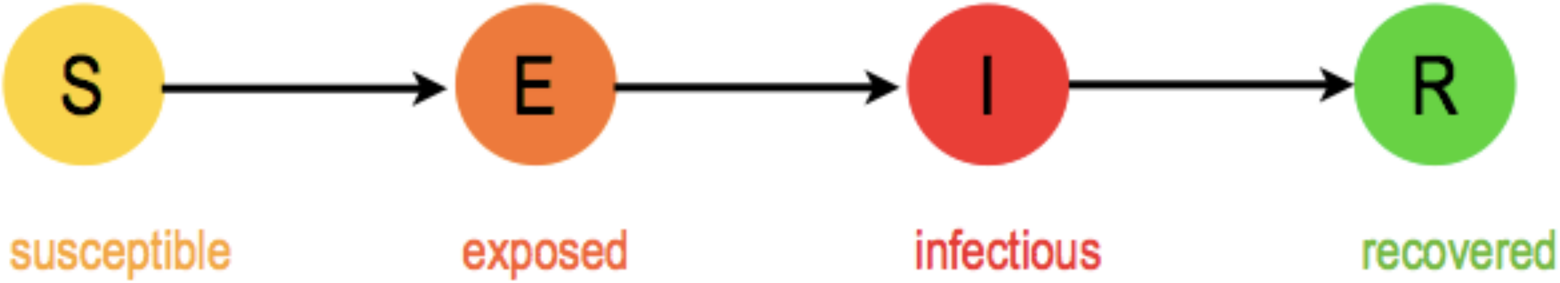
Homogeneous mixing: summary

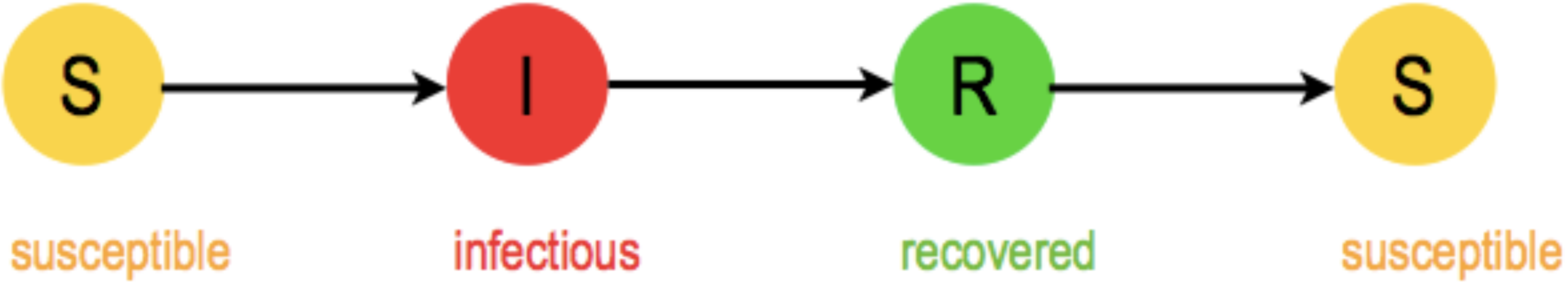


immunization threshold bringing the system under the epidemic threshold by depleting the susceptible population

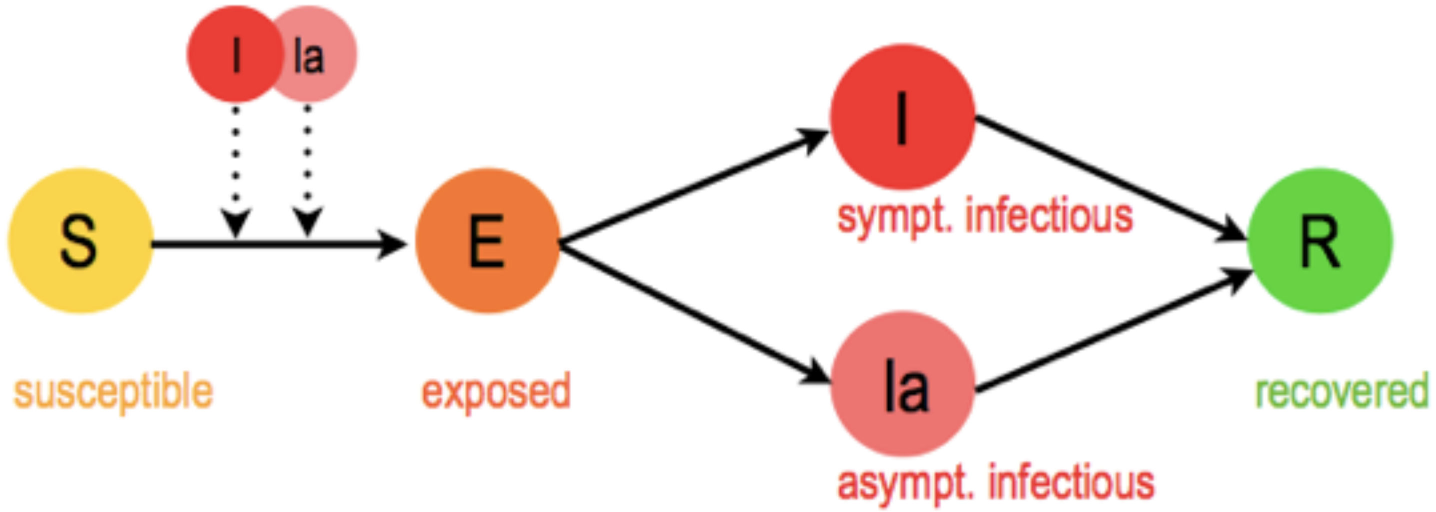
$$g_c = 1 - \mu/(\beta\langle k\rangle)$$

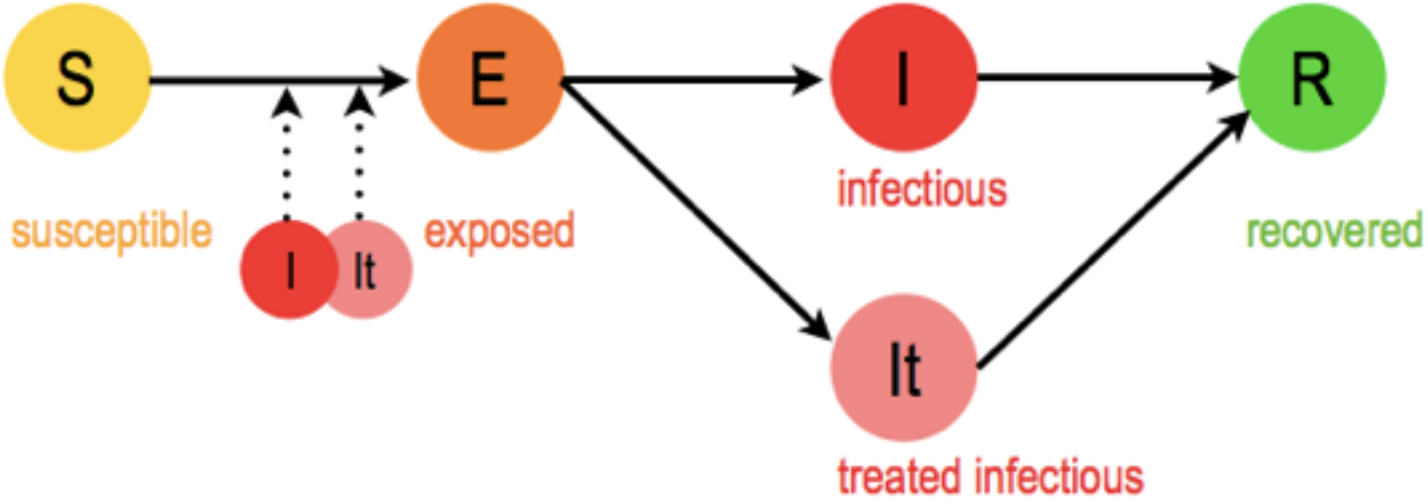
Going beyond: additional compartments





Going beyond: additional compartments





Going beyond: population structure

Different classes of individuals: age, gender, etc…

=> potentially different

- transmissibility
- contact rates

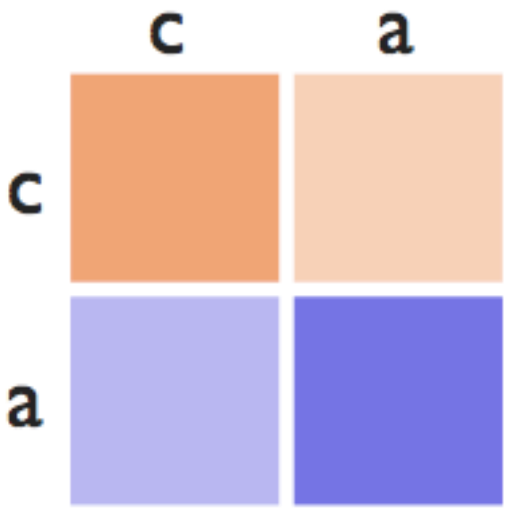
Going beyond: population structure

Different classes of individuals: age, gender, etc…
=> potentially different

- transmissibility
- contact rates

Contact matrices

Ex: flu => different contact rates for children and adults



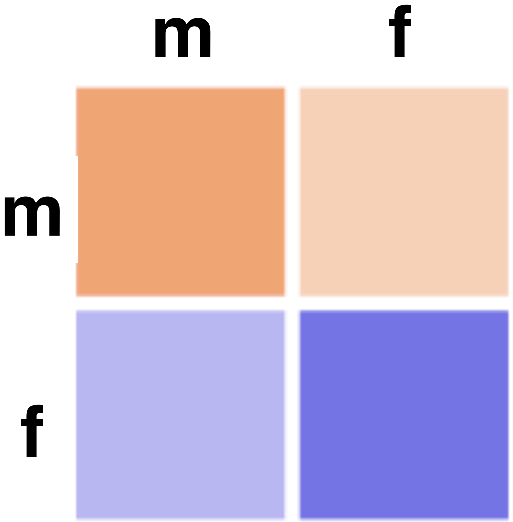
Going beyond: population structure

Different classes of individuals: age, gender, etc…
=> potentially different

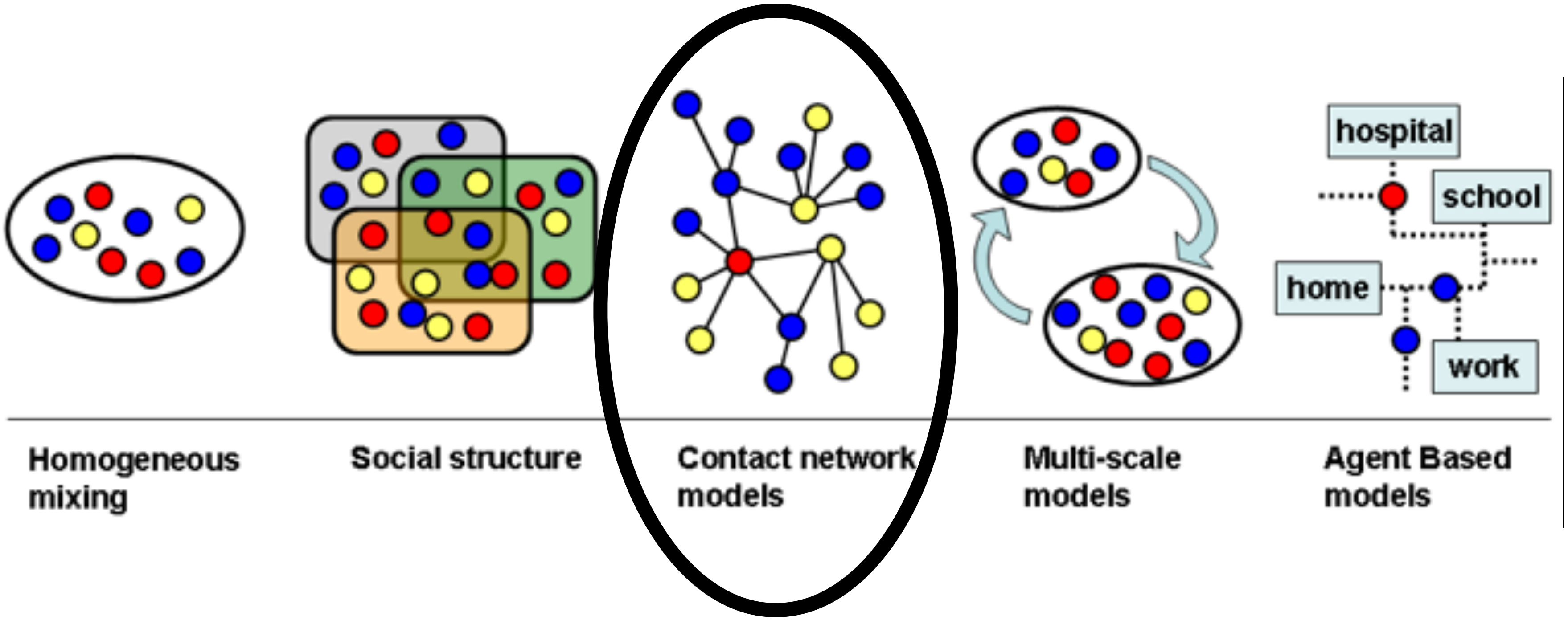
- transmissibility
- contact rates

→ Contact matrices

Ex: HIV => different transmissibility depending on gender



Wide spectrum of complications and complex features to include…

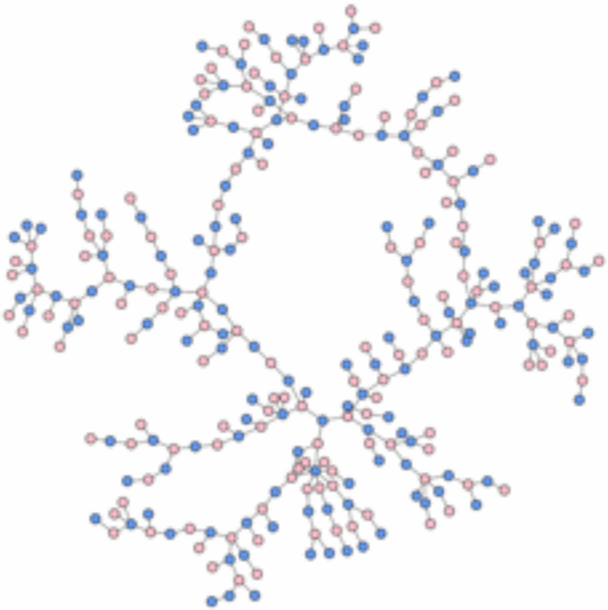


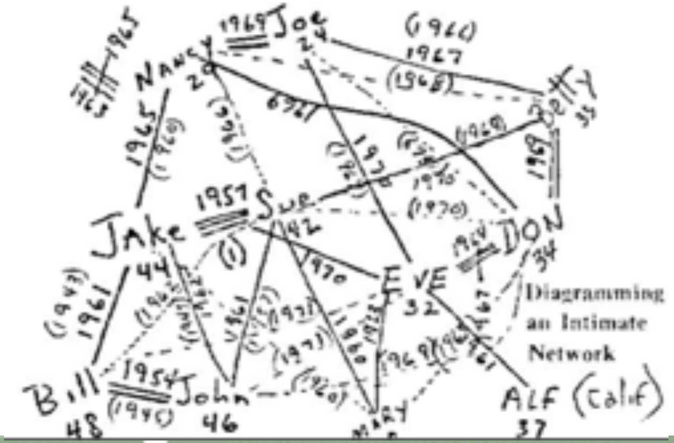
Simple → **Realistic**

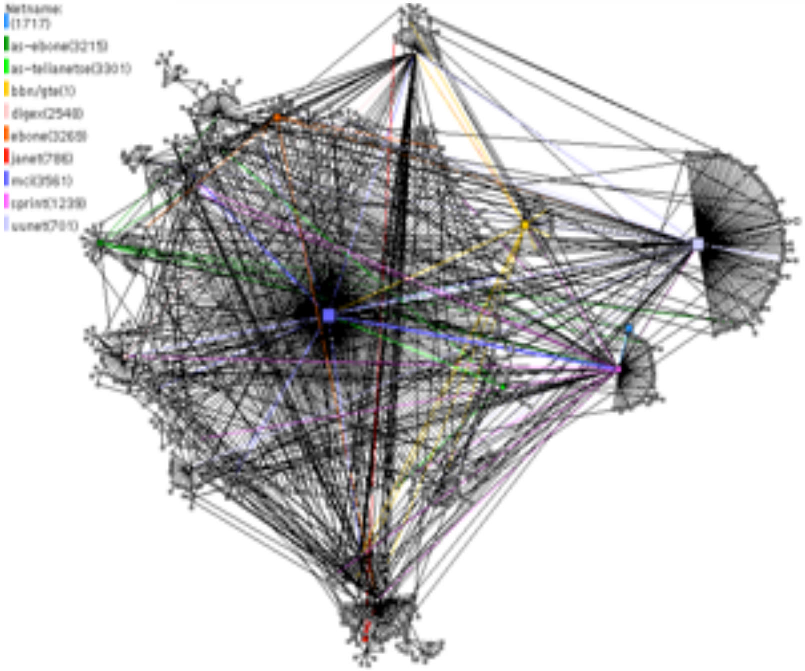
Ability to explain trends at a population level

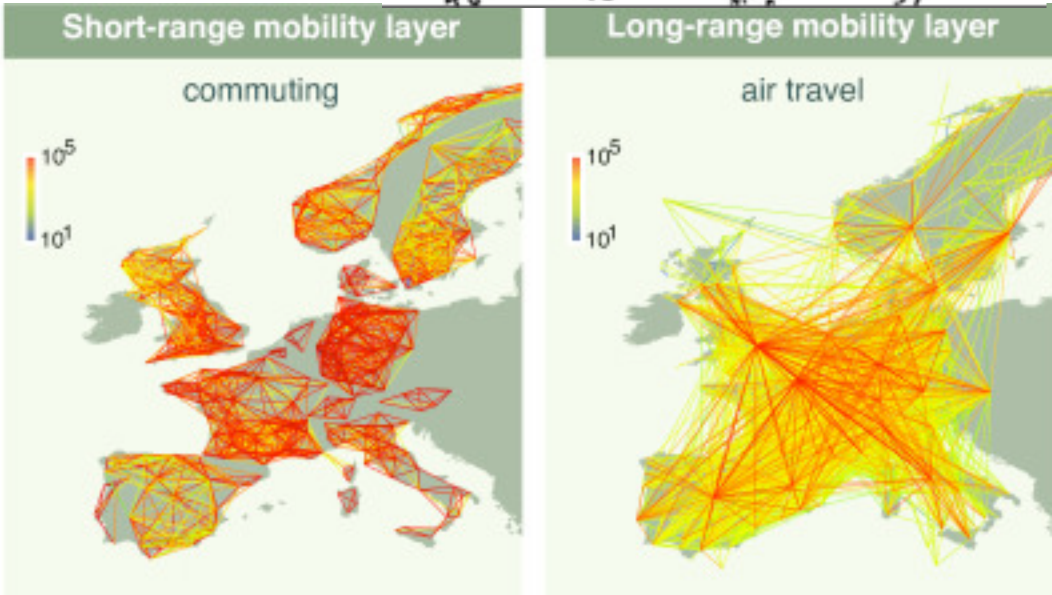
Model realism looses in transparency.
Validation is harder.

Complex networks







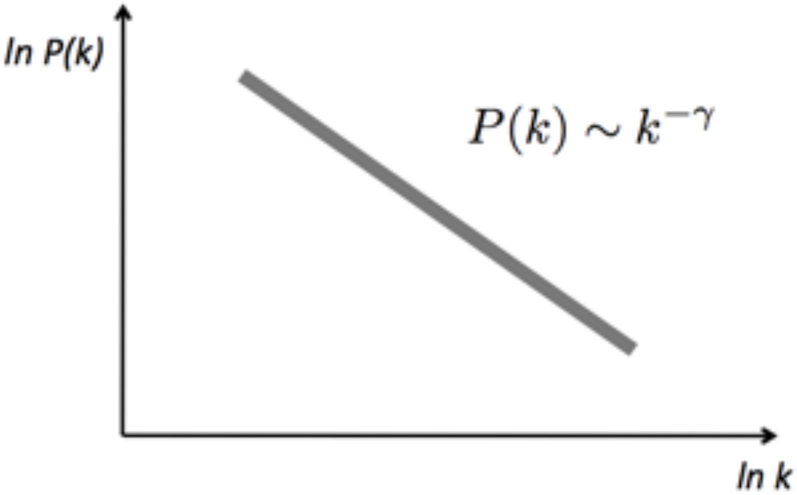


Complex networks

Diseases propagate on networks:

- Social (contact) networks
- Technological networks:
  - Internet, Web, P2P, e-mail...

...which are **complex, heterogeneous networks**



Usual mean-field: neglects the degree heterogeneity

Extension of mean-field theory to take it into account

Degree-based mean-field theory

Number of contacts (degree) can vary a lot

huge fluctuations ($\langle k^2 \rangle \gg \langle k \rangle$)

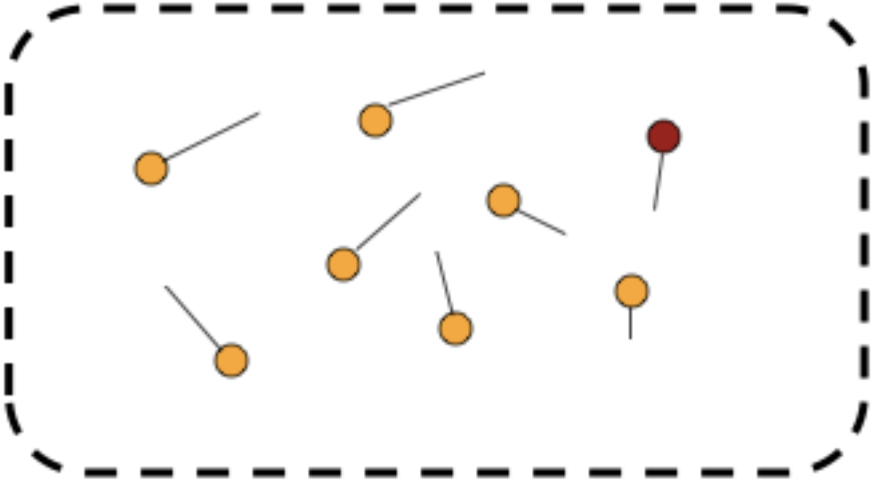
⬇

Heterogeneous (degree-based) mean-field: density of

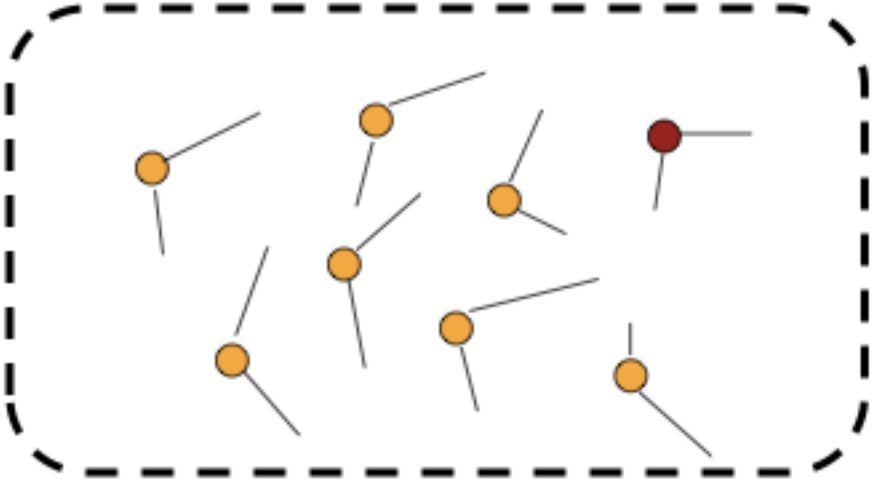
- Susceptible in the class of degree k, $s_k=S_k/N_k$
- Infectious in the class of degree k, $i_k=I_k/N_k$
- (Recovered in the class of degree k, $r_k=R_k/N_k$)

$$s(t)=\Sigma P(k)\, s_k,\ i(t)=\Sigma P(k)\, i_k,\ r(t)=\Sigma P(k)\, r_k$$

Degree-based representation



class of individuals with degree k=1

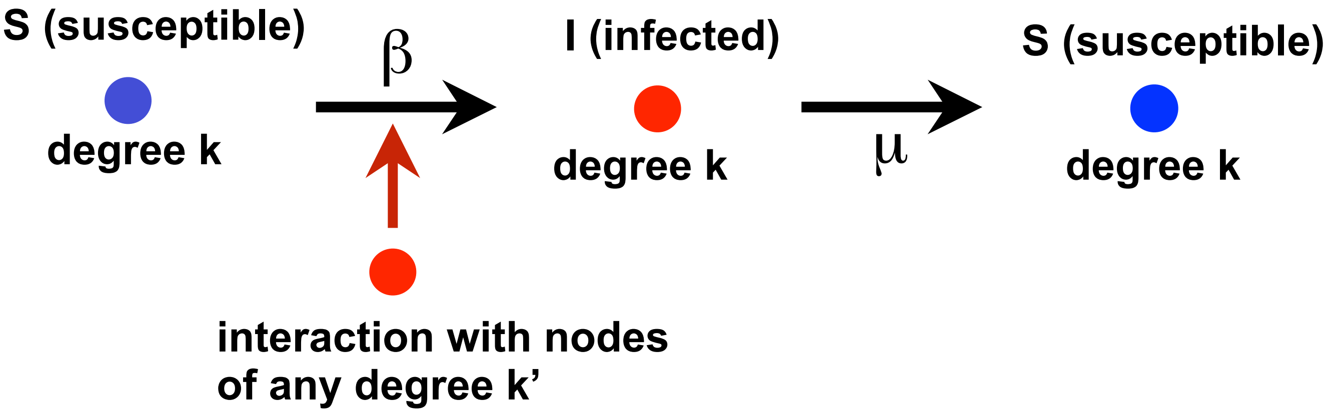


class of individuals with degree k=2

etc.

MF-like assumption: all individuals in a given class are “equivalent”

The SIS model in the degree-based MF theory



S (susceptible)

degree k

β

I (infected)

degree k

μ

S (susceptible)

degree k

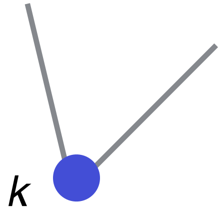
interaction with nodes of any degree k’

$$\frac{dI_k}{dt} = S_k(t) \times Proba(S_k \rightarrow I_k) - \mu I_k(t)$$

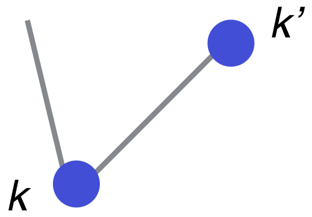
interaction with nodes of any degree k’

The SIS model in the degree-based MF theory

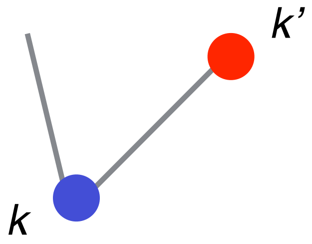
$$Proba(S_k \to I_k)$$



Number *k* of possible contacts



Proba of a contact with a node of degree *k’* $P(k'|k)$



Proba that the node is infectious: $\boxed{i_{k'}}$

Mean-Field

$P\,(k’|\,k)$ = the probability that a link originated in a node with connectivity *k* points to a node with connectivity *k’*

The SIS model in the degree-based MF theory

$$\frac{di_k}{dt} = \beta k(1 - i_k)\Theta_k - \mu i_k$$

Θ_k=Proba that any given link points to an infected node

$$\Theta_k = \sum_{k'} P(k'|k) i_{k'}$$

Mean-Field

$P(k'|k)$ = the probability that a link originated in a node with connectivity k points to a node with connectivity k'

The SIS model in the degree-based MF theory

$$\frac{di_k}{dt} = \beta k(1 - i_k)\Theta_k - \mu i_k \qquad \Theta_k = \sum_{k'} P(k'|k) i_{k'}$$

In uncorrelated networks: $\Theta_k = \Theta = \sum_{k'} \frac{k'}{\langle k \rangle} P(k') i_{k'}$

Short times, $i_k(t) << 1$

$$\boxed{\frac{d\Theta}{dt} = \left(\beta \frac{\langle k^2 \rangle}{\langle k \rangle} - \mu \right) \Theta}$$

Epidemic threshold condition $\boxed{\frac{\beta}{\mu} = \frac{\langle k \rangle}{\langle k^2 \rangle}}$

Long time limit, SIS model

$$\frac{di_k}{dt}=0 \Longrightarrow i_k(\infty)=\frac{\beta k\Theta_k(\infty)}{\beta k\Theta_k(\infty)+\mu}$$

$$\Theta_k=\sum_{k'}P(k'|k)i_{k'}$$

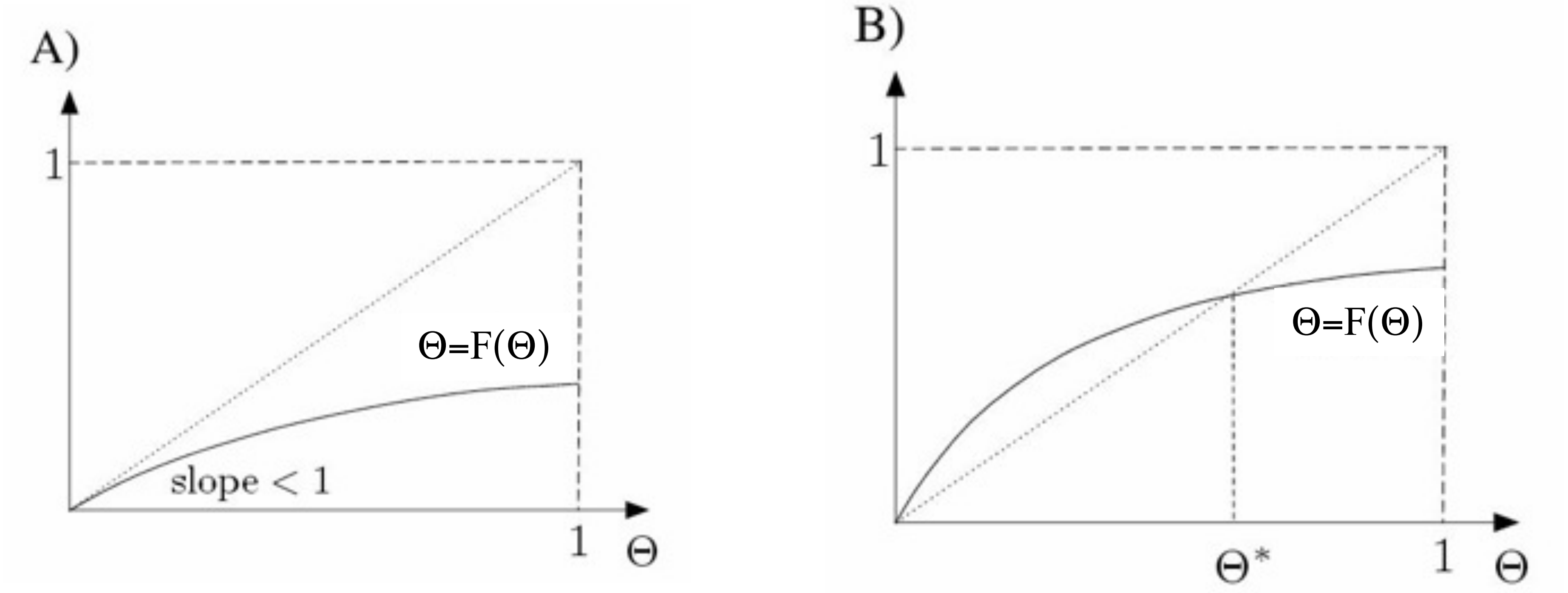
In uncorrelated networks: $\Theta_k=\Theta=\sum_{k'}\frac{k'}{\langle k\rangle}P(k')i_{k'}$

Sum over *k*

$$\Theta(\infty)=\frac{1}{\langle k\rangle}\sum_{k}\frac{\beta k^2P(k)\Theta(\infty)}{\beta k\Theta(\infty)+\mu}$$

Self-consistent equation of the form x=F(x)
with F(0)=0, F’ > 0, F’’ < 0

Graphical solution



Epidemic threshold:

existence of a non-zero solution for Θ ⇔ F'(0) > 1 :

$$\Leftrightarrow \quad \sum_k \frac{\beta k^2 P(k)}{\mu\langle k\rangle} > 1 \quad \Leftrightarrow \quad \frac{\beta\langle k^2\rangle}{\mu\langle k\rangle} > 1$$

The SIS model in the degree-based MF theory

Epidemic threshold in uncorrelated networks

$$\boxed{\frac{\beta}{\mu} = \frac{\langle k \rangle}{\langle k^2 \rangle}}$$

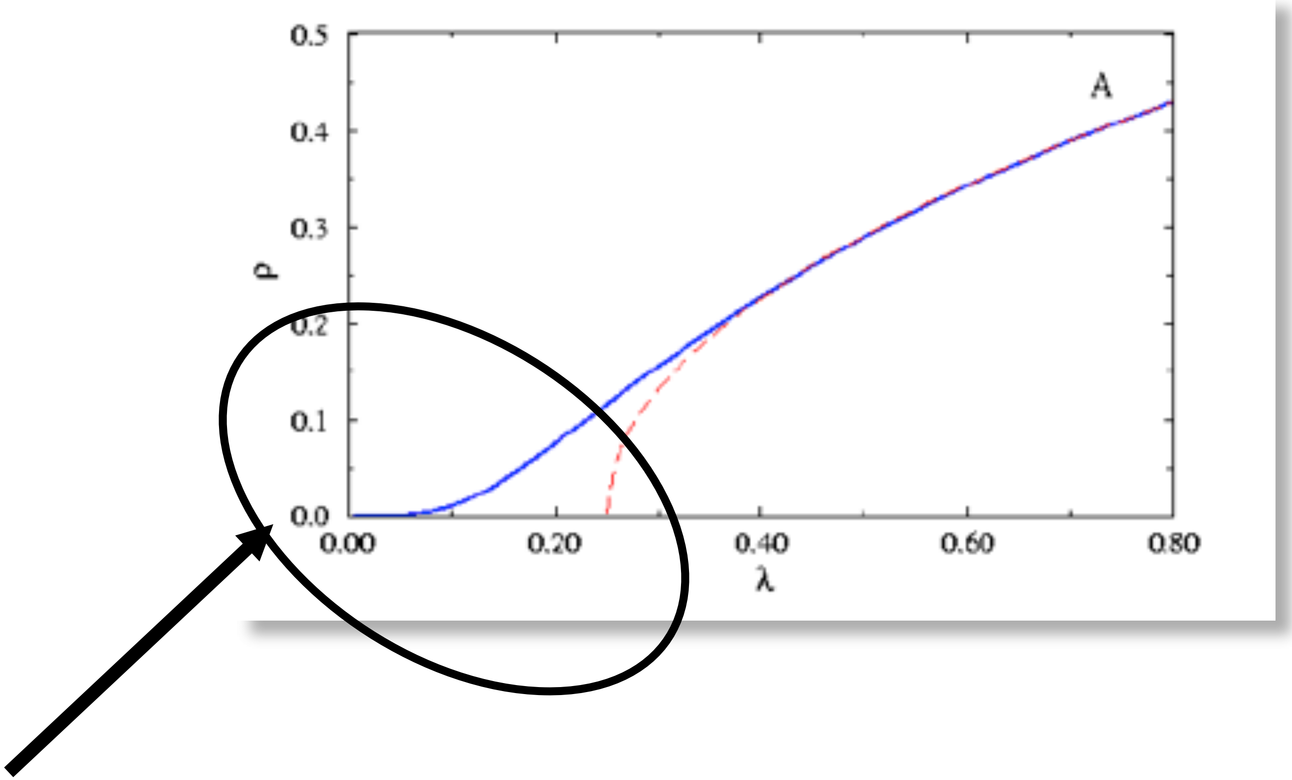
Heterogeneous, infinite network:

$$\langle k^2 \rangle \rightarrow \infty$$

Condition always satisfied

Finite prevalence for any spreading parameters

Epidemic phase diagram in heterogeneous networks



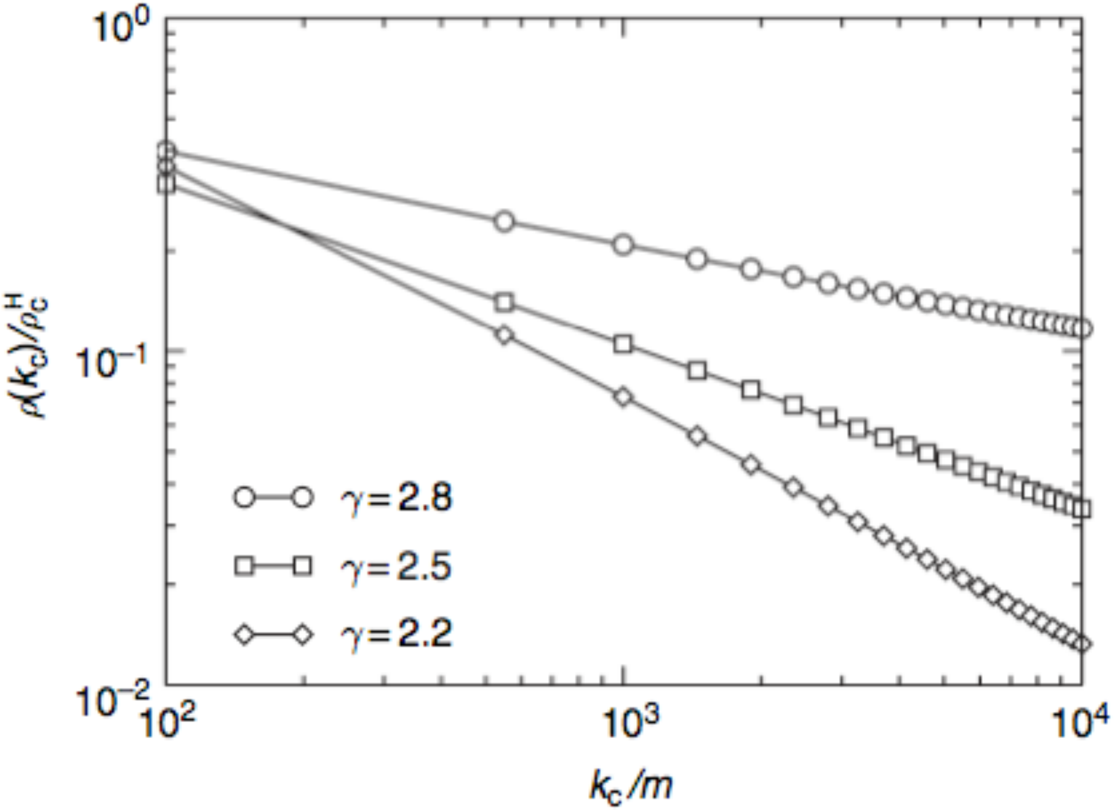
- Wide range of spreading rate with low prevalence
- Lack of healthy phase = standard immunization cannot drive the system below threshold!!!

Finite size effects

Finite number of nodes *N*

- ⇒ *Finite cut-off for P(k)*
- ⇒ *Finite* $\kappa = \langle k^2 \rangle / \langle k \rangle$
- ⇒ *Finite epidemic threshold*

Ratio of epidemic threshold to the value obtained in a homogeneous network:



Case of correlated networks

SIS model: $\frac{di_k}{dt} = \beta k(1-i_k)\Theta_k - \mu i_k \quad \Theta_k = \sum_{k'} P(k'|k) i_{k'}$

Short times:

$$\frac{di_k}{dt} \sim \sum_{k'} L_{kk'} i_{k'} \qquad L_{kk'} = -\mu\delta_{kk'} + \beta k P(k'|k)$$

Solution $i_k=0$ unstable iff there exists at least one positive eigenvalue

Λ_m largest eigenvalue of $C_{kk'} = kP(k'|k)$

Epidemic threshold: $\boxed{\frac{\beta}{\mu} = \frac{1}{\Lambda_m}}$

diverges in Markovian scale-free networks

Boguna, Pastor-Satorras, Vespignani, Phys. Rev. Lett. 90:028701 (2003), and arXiv:cond-mat/0301149

Spreading dynamics

Short times: $$\frac{d\Theta}{dt} = \left(\beta \frac{\langle k^2 \rangle}{\langle k \rangle} - \mu\right) \Theta$$

=> Exponential growth: $\Theta = \Theta_0 \exp(t/\tau)$

$$\frac{di_k}{dt} \approx \beta k \Theta - \mu i_k$$

$$i_k \propto k \exp(t/\tau)$$

Consequences on immunization strategies

Uniform immunization:

Fraction g of randomly chosen immunized (vaccined) individuals:

$$\beta \to \beta\,(1-g)$$

=> inefficient: need $(1-g)\frac{\beta}{\mu} < \frac{\langle k\rangle}{\langle k^2\rangle}$

$$\boxed{g > g_c = 1 - \frac{\mu}{\beta}\frac{\langle k\rangle}{\langle k^2\rangle}} \longrightarrow \text{tends to } 1$$

Proportional immunization

g_k fraction of immunized individuals of degree k, such that:

$$\beta k(1-g_k) = \beta' = cst$$

$$\frac{di_k}{dt} = \beta'(1-i_k)\Theta_k - \mu i_k$$

Short times (uncorr. nets):

$$\frac{d\Theta}{dt} = (\beta' - \mu)\,\Theta$$

Epidemic threshold recovered!

Efficient immunization: need $\beta' < \mu$ i.e., $\boxed{g_k > 1 - \frac{\mu}{\beta k}}$

Pastor-Satorras, Vespignani, Phys Rev E 65:036104 (2002)

Targeted immunization

=> immunize fraction g of individuals with largest connectivity

need: $$\frac{\langle k \rangle_g}{\langle k^2 \rangle_g} > \frac{\beta}{\mu}$$

similar to targeted attacks!!!

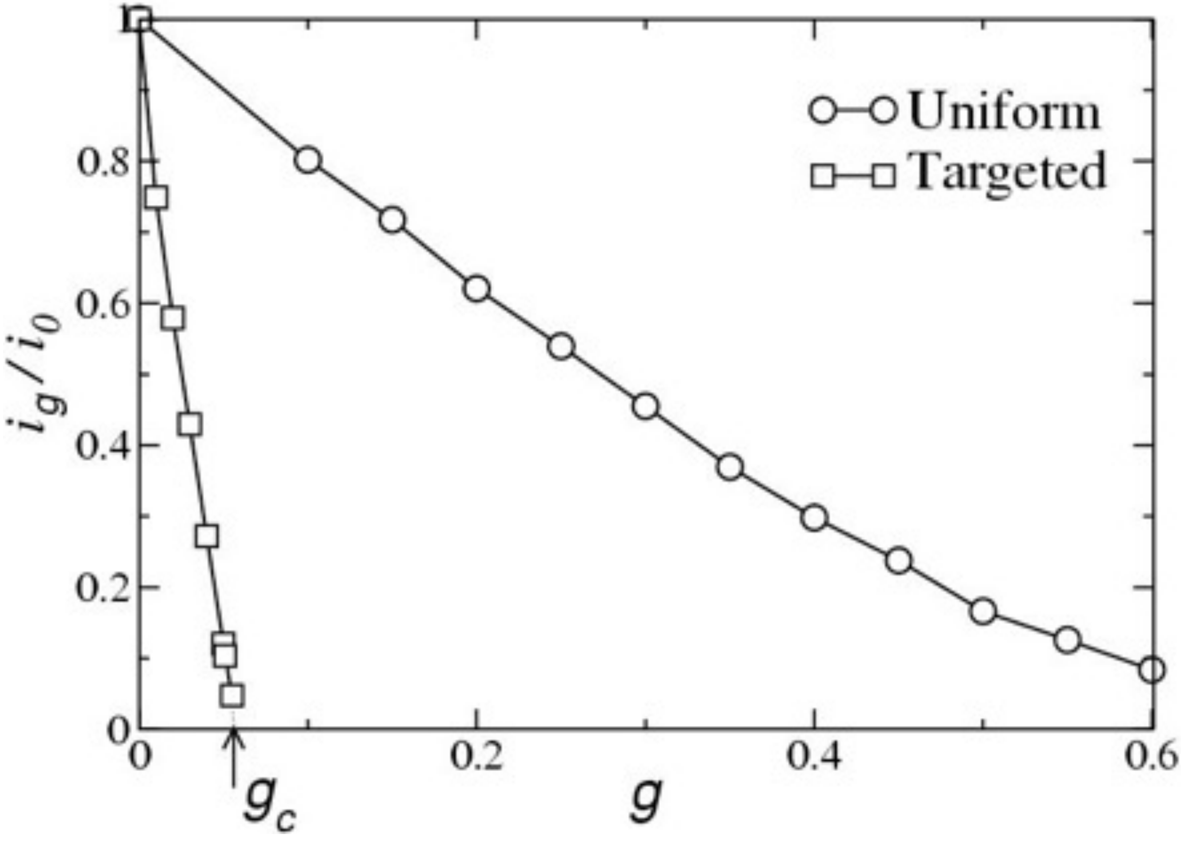
immunizing ⇔ removing nodes and links

Ex of explicit calculation for BA network:

$g_c \propto \exp(-2\mu/m\beta)$

Pastor-Satorras, Vespignani, Phys Rev E 65:036104 (2002)

Immunization



NB: when network's topology unknown: acquaintance immunization

[Cohen, Havlin, ben-Avraham, Phys Rev Lett 91:247901 (2003)]

What does HMF neglect

1. Structural correlations in the network

(HMF equivalent to an annealed network approximation)

2. Dynamical correlations

(emerging during the spreading process)

Beyond HMF: Quenched mean-field

Chakrabarti et al., Epidemic Thresholds in Real Networks, ACM Trans. Inf. Syst. Secur. 10, 1 (2008)

Gomez et al., Discrete-time Markov chain approach to contact-based disease spreading in complex networks, EPL 89 38009 (2010)
P. Van Mieghem, arXiv:1402.1731

"Quenched"/"Individual-based" mean field theory:

- write the evolution equation for the probability that a node i is infected
- take into account the real connections of the network as given by the adjacency matrix
- neglect correlations (hence mean-field)
- check for the stability of the absorbing state of zero infection

Quenched MF for the SIS model

P. Van Mieghem, arXiv:1402.1731

$X_i(t)$ = 0 for S nodes, 1 for I nodes

$E[X_i(t)] = \rho_i^I(t)$ = probability that i is infected

Real network structure

$$\frac{dE[X_i(t)]}{dt} = E\left[-\mu X_i(t) + \beta(1 - X_i(t)) \textstyle\sum_{j=1}^{N} a_{ij} X_j(t)\right]$$

$\lambda = \beta/\mu$

$$\frac{d\rho_i^I(t)}{dt} = -\rho_i^I(t) + \lambda \sum_{j=1}^{N} a_{ij}\rho_j^I(t) - \lambda \sum_{j=1}^{N} a_{ij} E\left[X_i(t)X_j(t)\right]$$

Mean-Field

$$\frac{d\rho_i^I(t)}{dt} = -\rho_i^I(t) + \lambda(1 - \rho_i^I(t)) \sum_{j=1}^{N} a_{ij}\rho_j^I(t)$$

MF: correlations neglected

Quenched MF for the SIS model

$$\frac{d\rho_i^I(t)}{dt} = -\rho_i^I(t) + \lambda(1 - \rho_i^I(t)) \sum_{j=1}^{N} a_{ij}\rho_j^I(t)$$

Linear approximation for the stability of the epidemic-free state

$$\frac{d\vec{\rho^I}}{dt} = -\vec{\rho^I} + \lambda A \vec{\rho^I}$$ A = adjacency matrix

⇒ $$\vec{\rho^I} = \sum_{\alpha} \vec{\rho_\alpha} \exp\left((\lambda \nu_\alpha - 1)t\right)$$

ρ_α, ν_α eigenvalues and eigenvectors of the adjacency matrix A

⇒ $\vec{\rho^I} = 0$ unstable iff the largest eigenvalue of A is larger than 1/λ

Quenched MF for the SIS model

=> epidemic threshold for SIS given by

$$\lambda_c = 1/\Lambda_m$$

where Λ_m is the largest eigenvalue of the adjacency matrix

Quenched vs Degree-based MF

Use ***annealed*** adjacency matrix in quenched MF equation $a_{ij} \equiv a_{k_i k_j}$

Ex for uncorrelated networks $a_{k_i k_j} = \dfrac{k_i k_j}{N\langle k\rangle}$

Sum equation $$\frac{d\rho_i^I(t)}{dt} = -\rho_i^I(t) + \lambda(1-\rho_i^I(t))\sum_{j=1}^{N} a_{ij}\rho_j^I(t)$$

over all nodes of degree k:

recover heterogeneous mean-field equation for $\rho_k = \dfrac{1}{N_k}\displaystyle\sum_{i|k_i=k} \rho_i^I$

Quenched MF for the SIS model

For random scale-free networks, it is possible to obtain the scaling of Λ_m (Chung, Lu, Vu, Proc. Natl. Acad. Sci. USA 100, 6313 (2003))

=> epidemic threshold for SIS given by

$$\lambda_c = \begin{cases} 1/\sqrt{k_{max}} \text{ if } \gamma > 5/2 \\ \langle k \rangle / \langle k^2 \rangle \text{ if } 2 < \gamma < 5/2 \end{cases}$$

=> epidemic threshold vanishes in the thermodynamic limit in power-law distributed networks

- for any value of γ, even larger than 3,
- as long as k_{max} is a diverging function of the network size N

=> role of the hubs

Numerics

Kitsak et al., Identification of influential spreaders in complex networks. Nat. Phys. 6, 888–893 (2010)

Numerical study => most efficient spreaders are located at the innermost, dense core of the network, as identified by means of a k-core decomposition

=> apparent contradiction?

Numerics

Castellano & Pastor-Satorras, Competing activation mechanisms in epidemics on networks, Scientific Reports 2, 371 (2012)

=> numerical investigation, measure of:

- density of infected vertices in the whole network
- density of infected when the dynamics takes place (in isolation) on the **k-core of highest index** (maximum k-core)
- density of infected when the dynamics takes place (in isolation) on the star-graph centered around the **hub** of the network, with largest degree

Beyond HMF

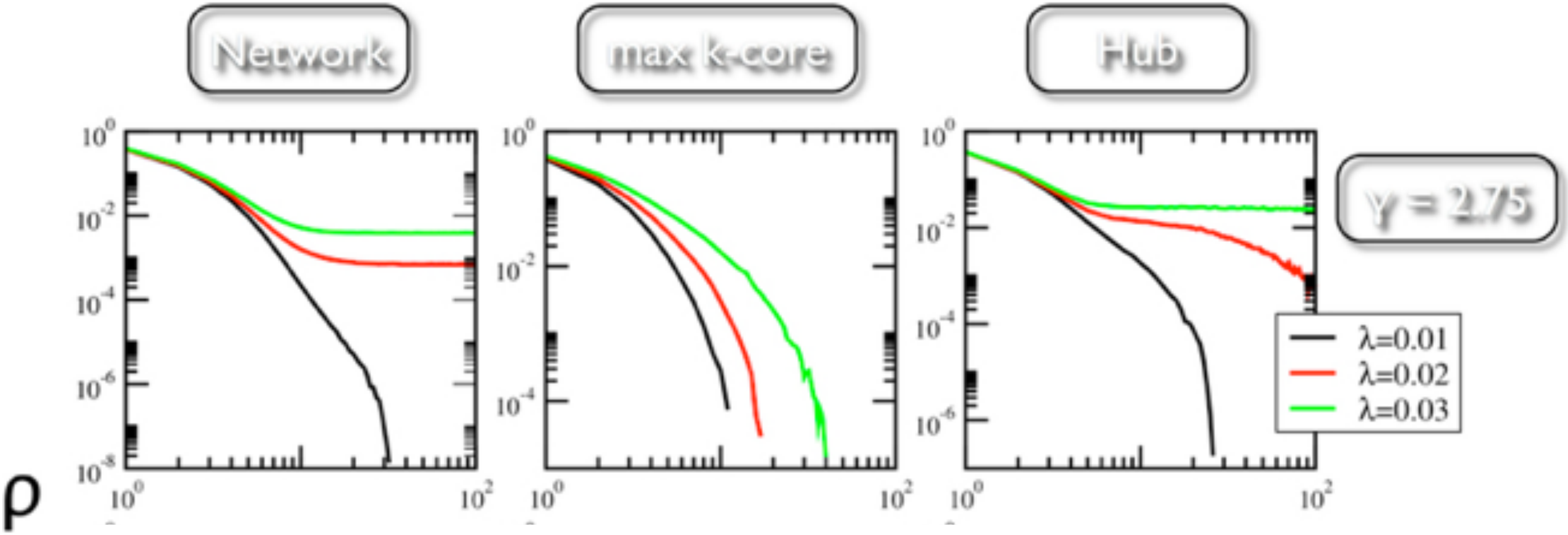
Castellano & Pastor-Satorras, Competing activation mechanisms in epidemics on networks, Scientific Reports 2, 371 (2012)

Results, as β increases:

large γ: onset of active stationary state for the whole network for values of β at which the star graph around the hub starts being active, while the max-k-core is subcritical

Numerics

Castellano & Pastor-Satorras, Competing activation mechanisms in epidemics on networks, Scientific Reports 2, 371 (2012)



Beyond HMF

Castellano & Pastor-Satorras, Competing activation mechanisms in epidemics on networks, Scientific Reports 2, 371 (2012)

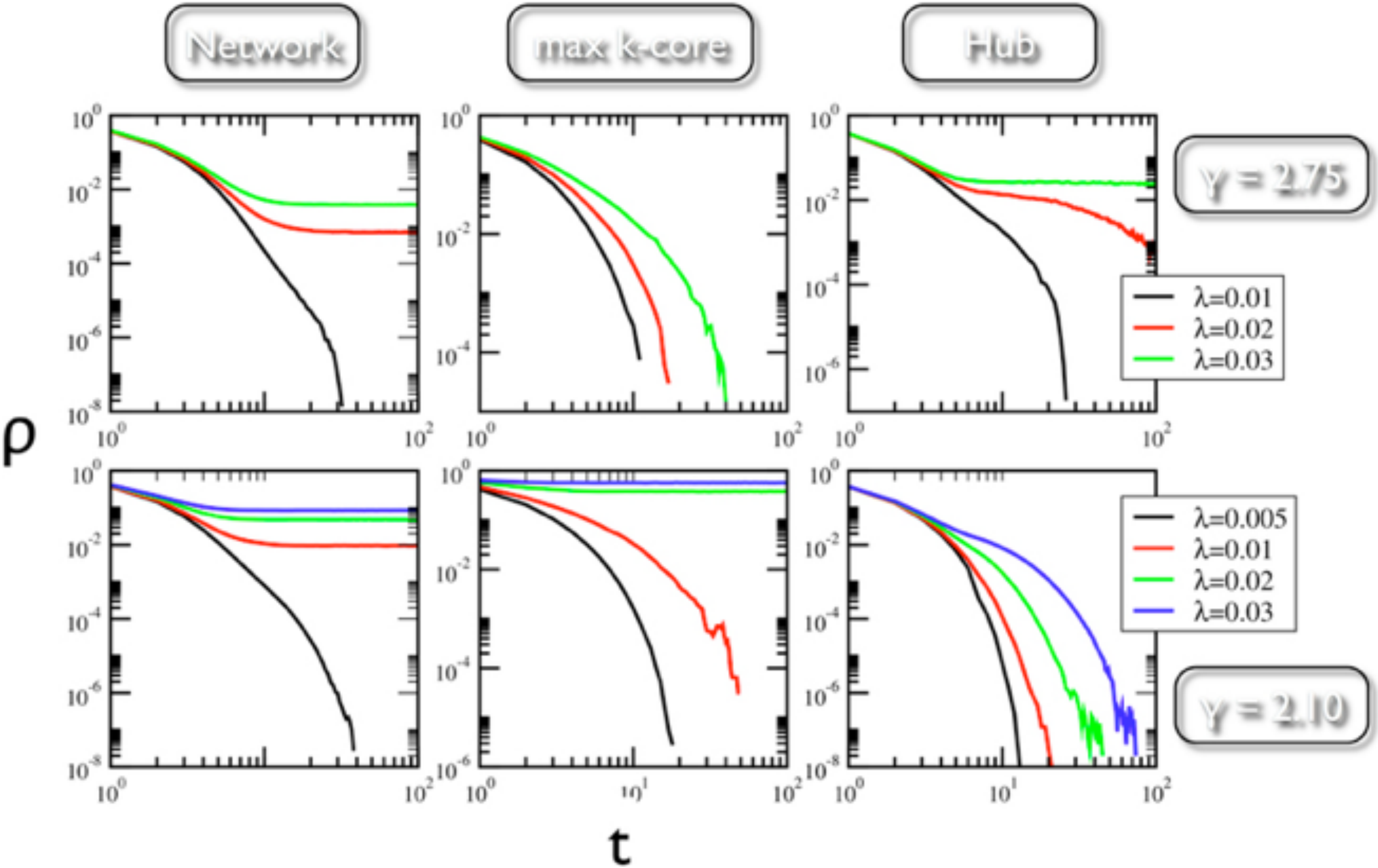
Results, as β increases:

large γ: onset of active stationary state for the whole network for values of β at which the star graph around the hub starts being active, while the max-k-core is subcritical

small γ: onset of active stationary state for the whole network for values of β at which the max k-core starts being active, while the star graph is subcritical

Numerics

Castellano & Pastor-Satorras, Competing activation mechanisms in epidemics on networks, Scientific Reports 2, 371 (2012)



Numerics

Castellano & Pastor-Satorras, Competing activation mechanisms in epidemics on networks, Scientific Reports 2, 371 (2012)

- the max-k-core is a homogeneous network, hence has an epidemic threshold $\lambda_S \sim 1/k_S$
- the scaling of k_S is known for scale-free networks: $k_S \sim k_{max}^{\gamma-3}$ (same scaling as $\langle k^2 \rangle$)
- for γ < 5/2, this gives back the HMF result
- for γ > 5/2, the hub is responsible

“When γ < 5/2, the epidemic transition is collectively triggered by the vertices in the innermost core and the threshold is correspondingly given by $1/\langle k^2 \rangle$, as in HMF theory. On the other hand, for γ > 5/2, the hub triggers the global activity, and the threshold is given by $1/\sqrt{k_{max}}$ “

Numerics

Castellano & Pastor-Satorras, Competing activation mechanisms in epidemics on networks, Scientific Reports 2, 371 (2012)

Ferreira, Castellano, Pastor-Satorras, arXiv:1206.6728, Phys. Rev. E 86, 041125 (2012)

=> for exponents < 5/2, HMF substantially correct

=> for exponents > 5/2, QMF better than HMF, but dynamic correlations come into play => QMF still needs improvement

Some rigorous results

Chatterjee, Durrett, Contact processes on random graphs with power law degree distributions have critical value 0. Annals of Probability 37, 2332–2356 (2009).

Durrett, Some features of the spread of epidemics and information on a random graph, Proc. Natl. Acad. Sci. USA 107, 4491 (2010)

Mountford et al., Exponential extinction time of the contact process on finite graphs, arXiv:1203.2972

Van Mieghem, Exact Markovian SIR and SIS epidemics on networks and an upper bound for the epidemic threshold, arXiv:1402.1731

=> special cases, bounds and asymptotic exact results

Some more complications

Degree correlations

Clustering

Directed networks

Weights

DBMF: $\lambda_{kk'} \propto (kk')^{\sigma}$

IBMF: largest vp of $\Omega_{ij} = w_{ij} a_{ij}$

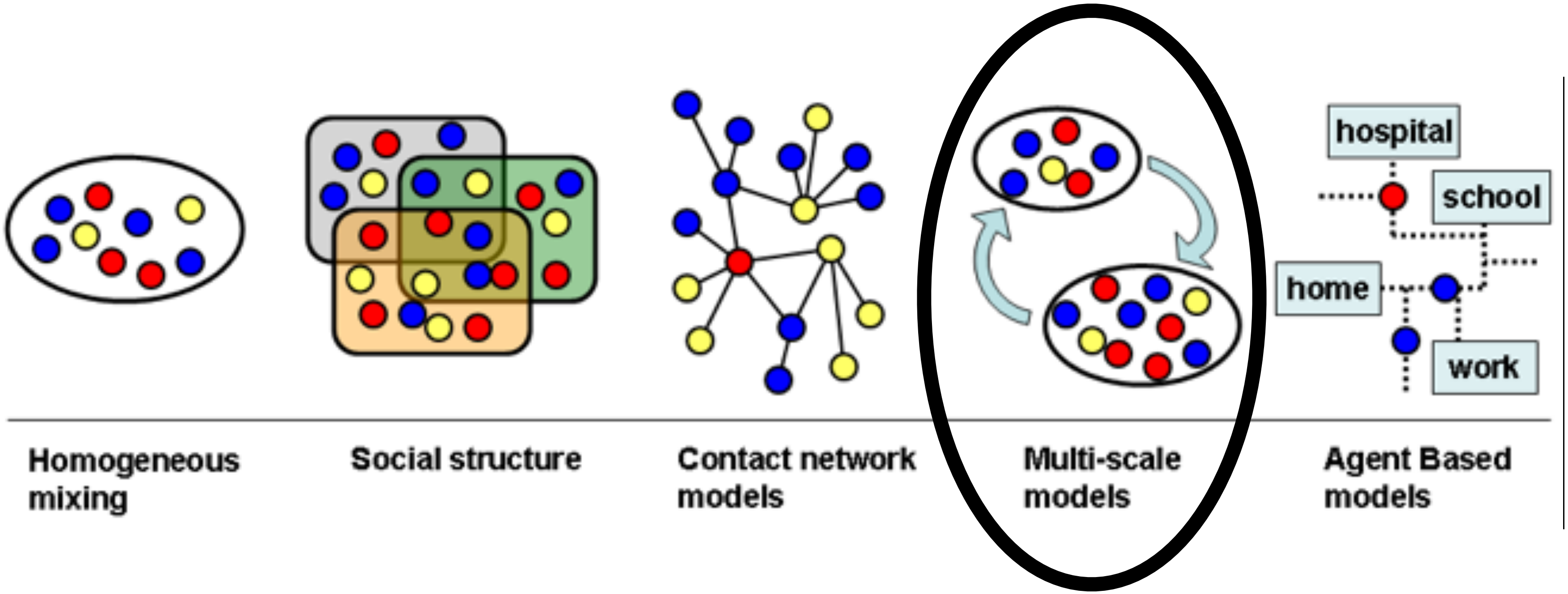
Community structures

Initial (local) faster spread, slowing down at global scale

Strength of weak ties (Granovetter 1973, Onnela et al. 2007)

Immunization of bridges

Wide spectrum of complications and complex features to include…



Simple → **Realistic**

Ability to explain trends at a population level

Model realism looses in transparency.
Validation is harder.

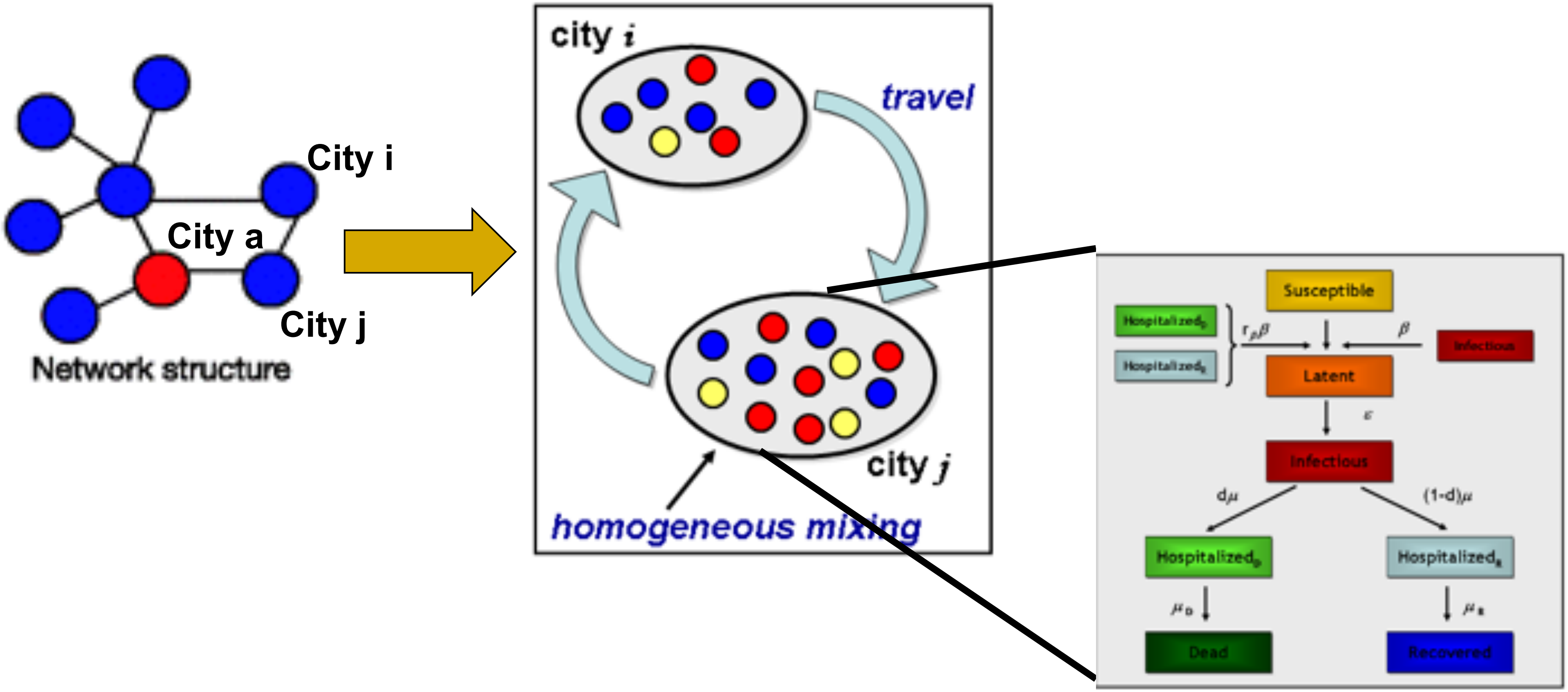
General framework:
reaction-diffusion processes

- Previous cases: (at most) *one* particle/individual *per site*
- In general: reaction-diffusion processes on networks
  => no restriction on the number of particles per site

“Particles”

- diffusing along edges
- reacting in the nodes

Meta-population models



Intra-population infection dynamics by stochastic compartmental modeling

Inside each population: homogeneous mixing

Baroyan *et al.* (1969)
Ravchev, Longini (1985)

Modeling of global epidemics propagation

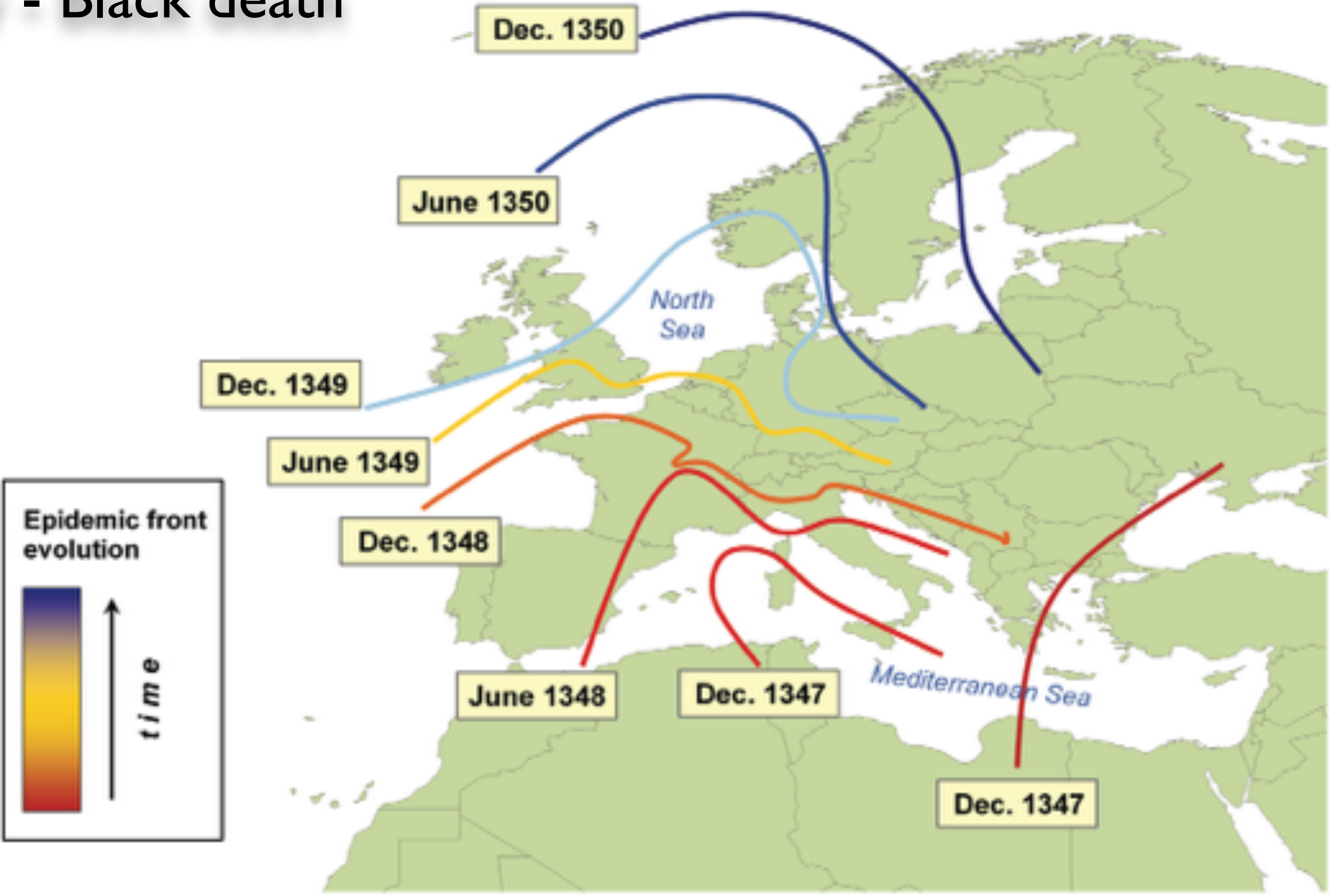
multi-level description :

- intra-city **epidemics**
- inter-city **travel**

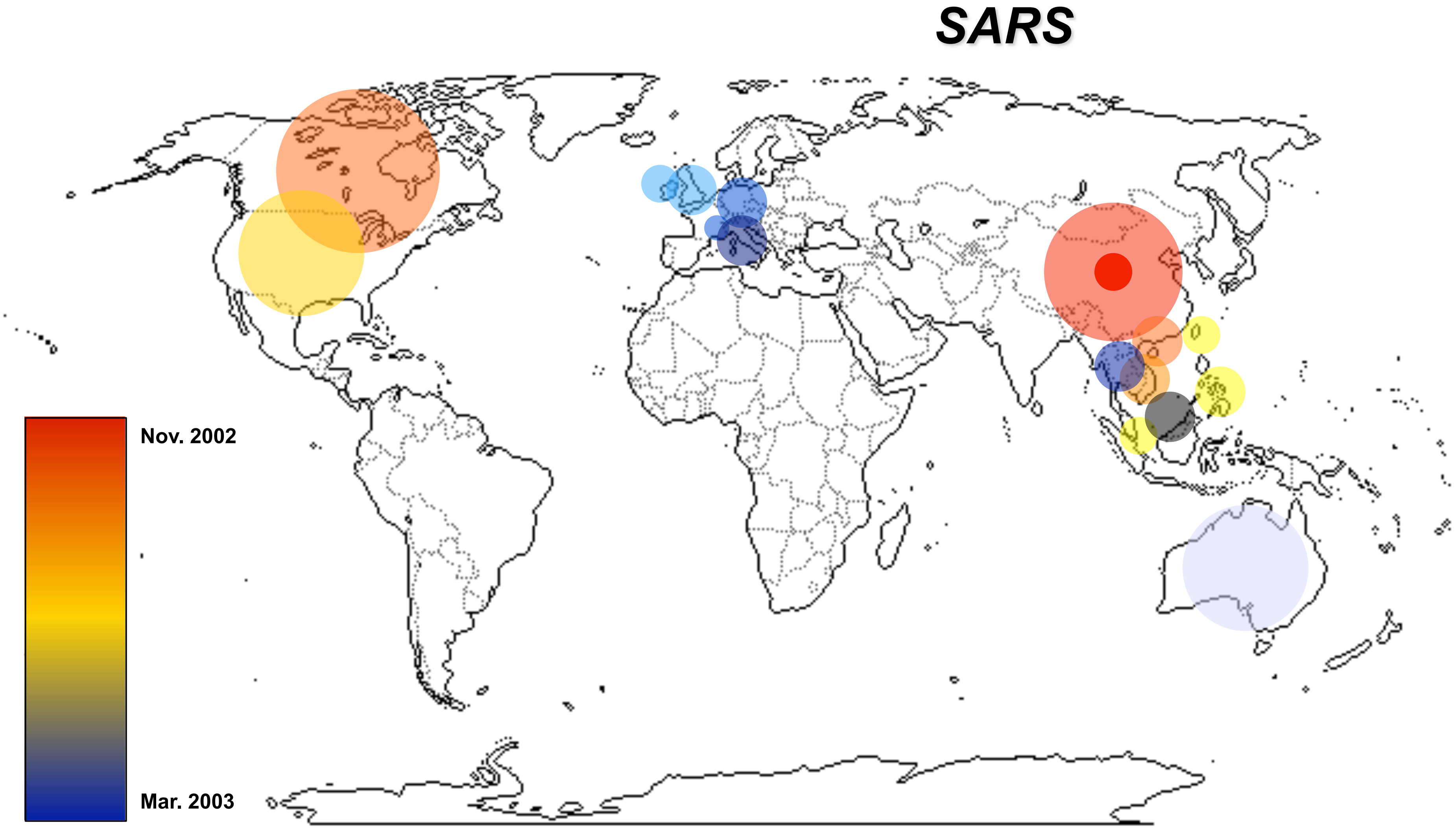
Baroyan *et al.* (1969)
Ravchev, Longini (1985)

Why is a large-scale approach needed?

14th century - Black death

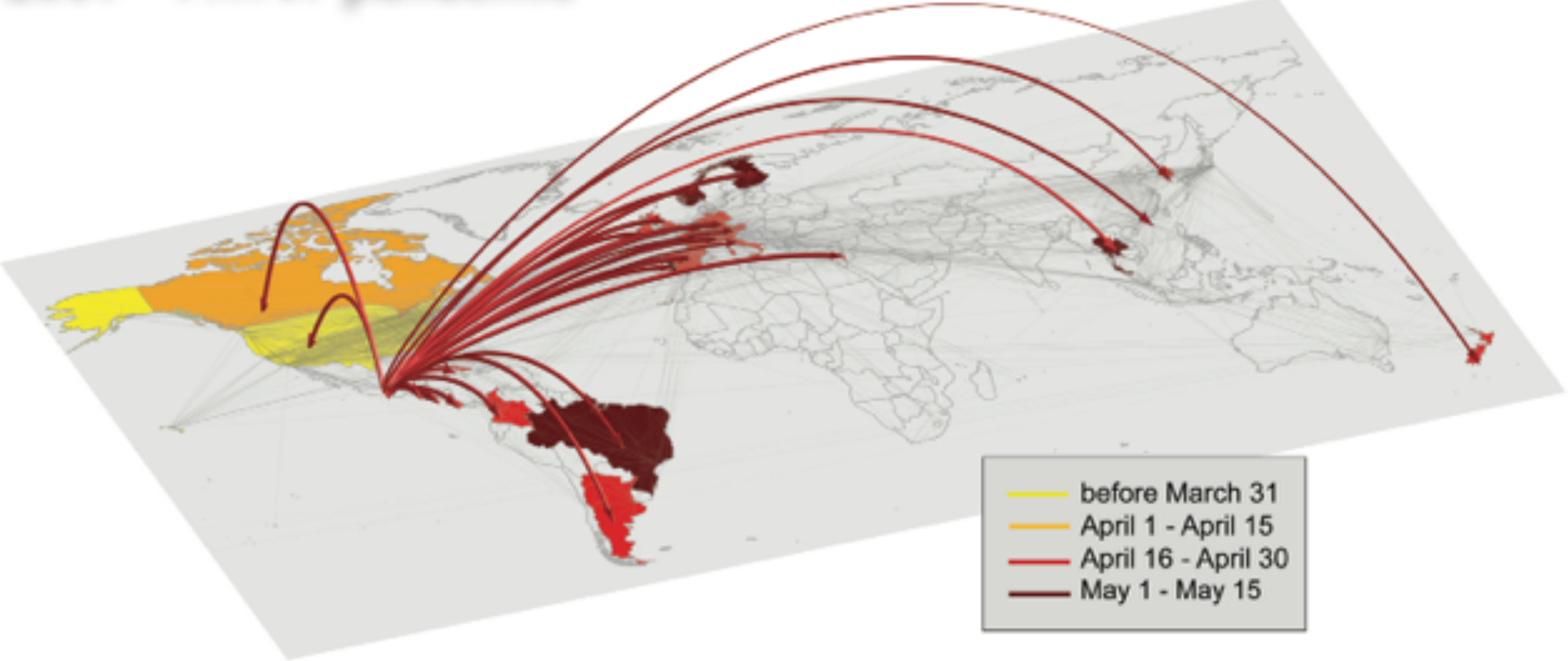


Why is a large-scale approach needed?



Why is a large-scale approach needed?

2009 - H1N1 pandemic

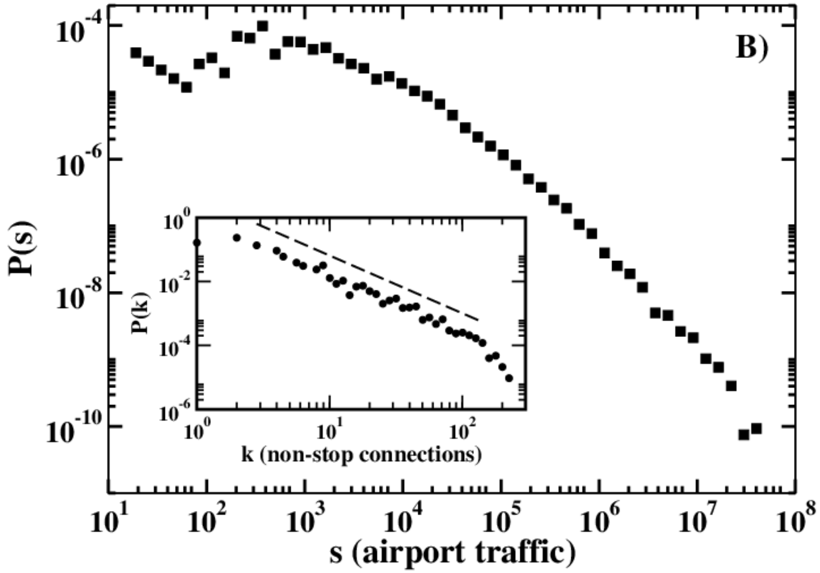


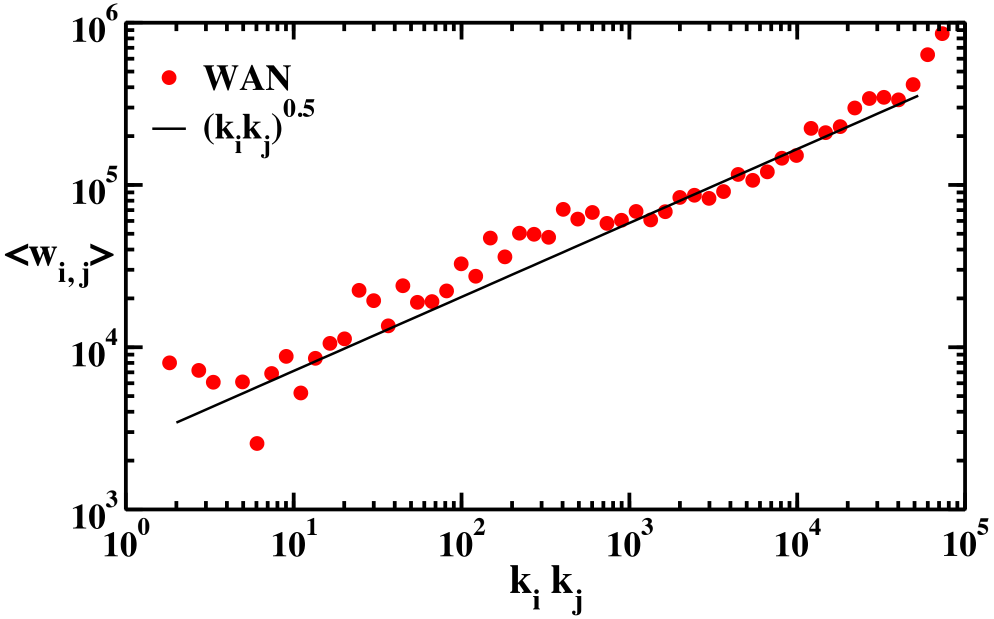
Bajardi et al, PLoS ONE (2011)

What has changed: availability of unprecedented amounts of data…..

- Transportation infrastructures
- Behavioral Networks
- Census data
- Commuting/traveling patterns

  - Different scales:
    - International
    - Intra-nation (county/city/municipality)
    - Intra-city (workplace/daily commuters/individuals behavior)





Barrat et al., PNAS (2004), Colizza et al. PNAS (2006)

A recent large-scale platform

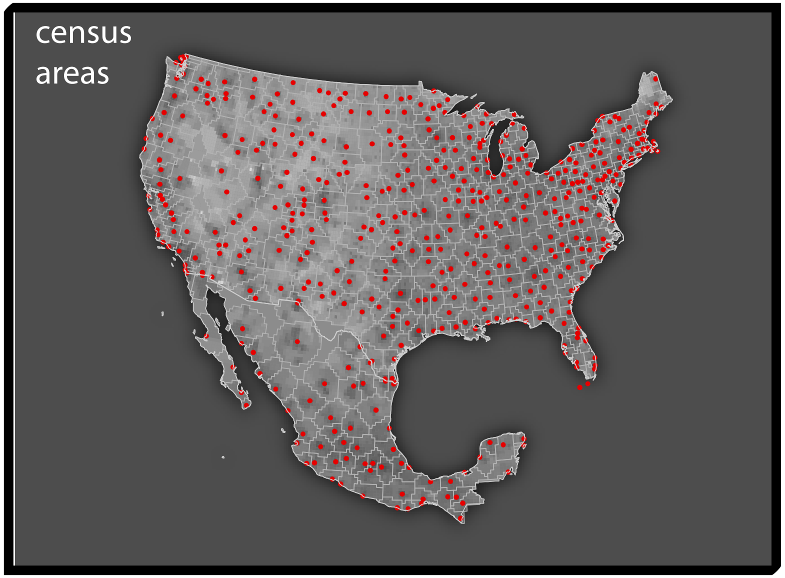
http://www.gleamviz.org

simulation platform for the worldwide propagation of diseases, used in real time during the H1N1 pandemic

D. Balcan, V. Colizza, B. Gonçalves, H. Hu, J.J. Ramasco, A. Vespignani

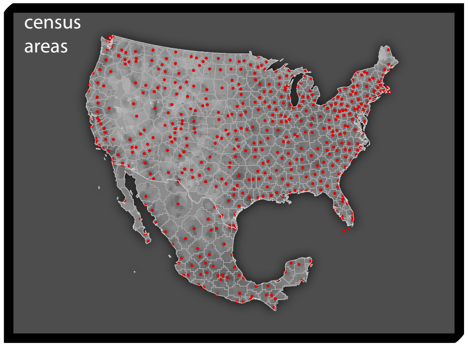
Proc. Natl. Acad. Sci. USA 106, 21484-21489 (2009)

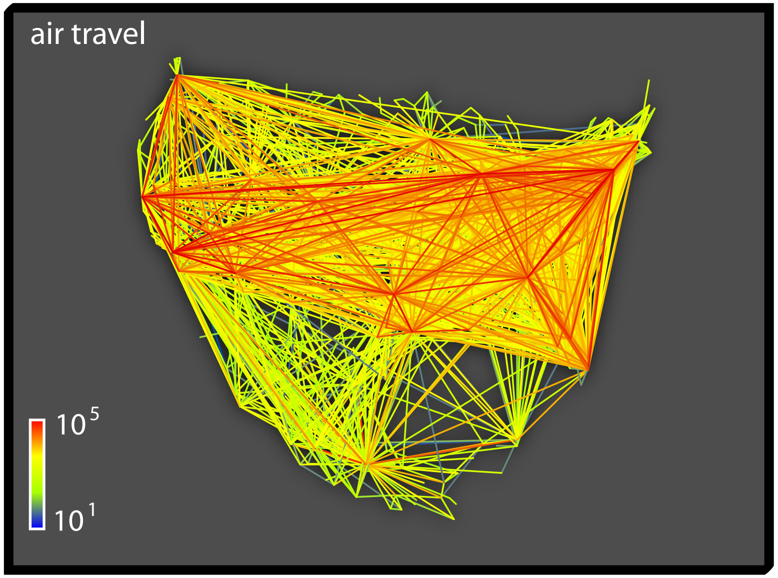
GLEaM in brief



Population distribution:
detailed population data from
1/4x1/4 degree tasselation.

GLEaM in brief

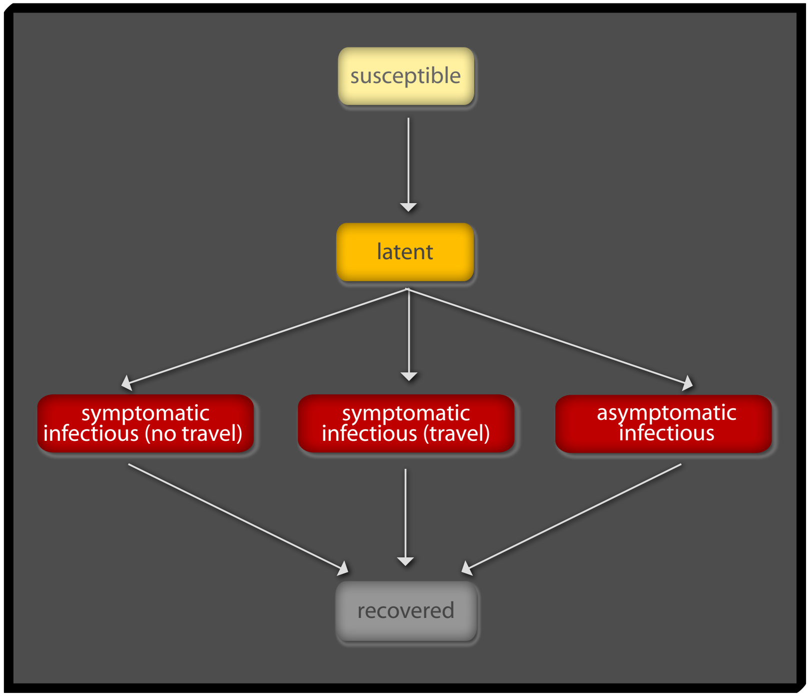




Local mobility:
census data from about 30 countries in the 5 continents extended to all countries.

Long range travel:
3362 cities in 220 countries.
More than 16000 connections with travel flows.

GLEaM in brief



susceptible

latent

symptomatic infectious (no travel)

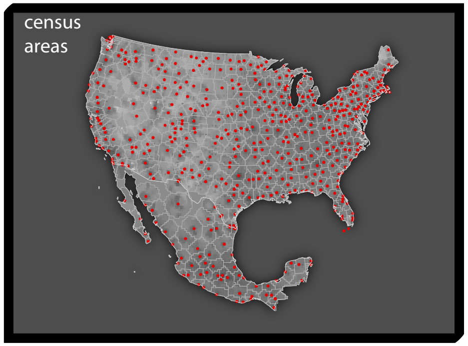
symptomatic infectious (travel)

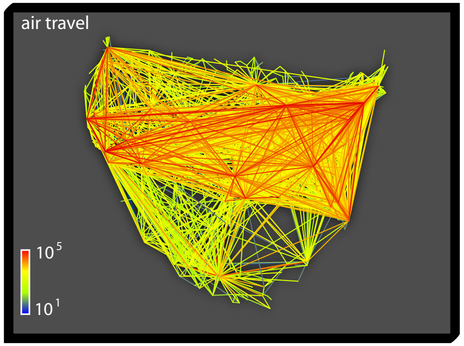
asymptomatic infectious

recovered

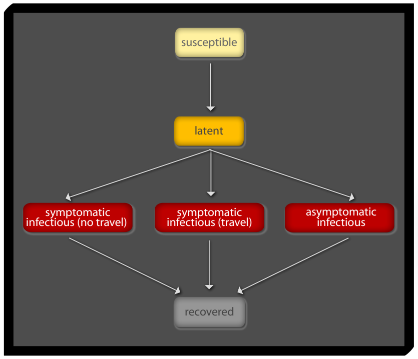
Epidemic compartmental model

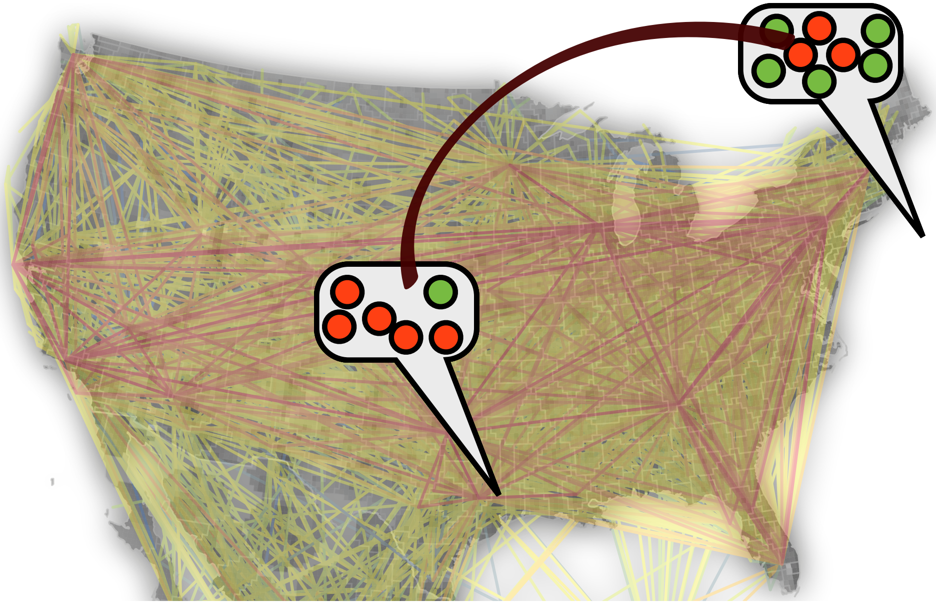
Metapopulation model with homogeneous mixing assumption.

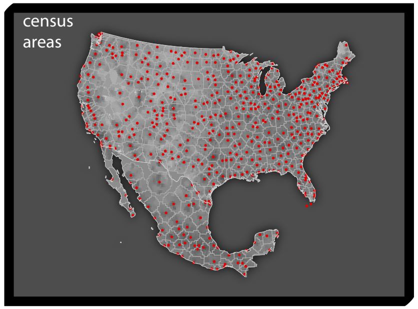


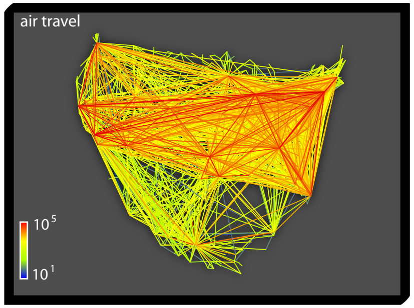


GLEaM in brief



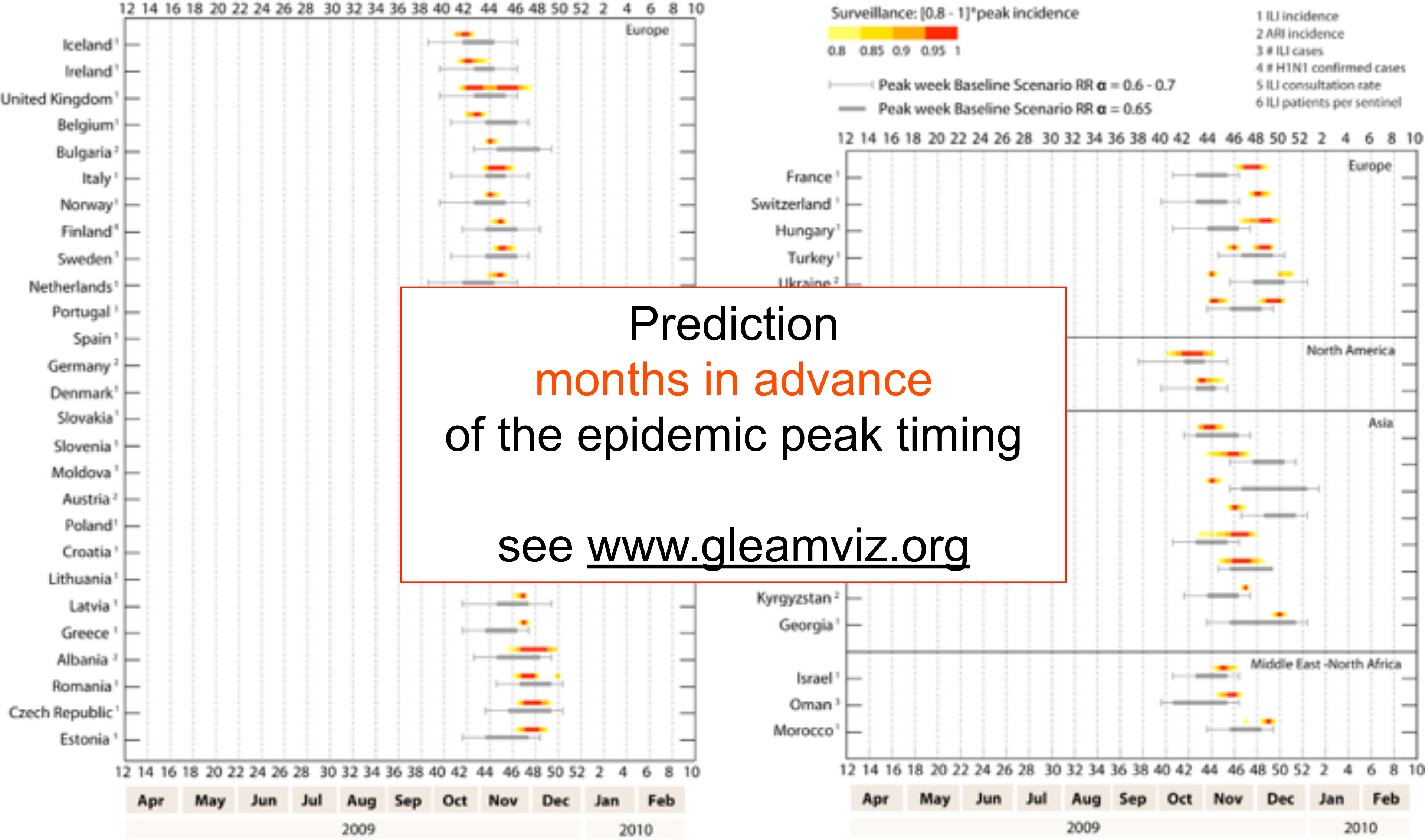






http://www.gleamviz.org

comparing with data



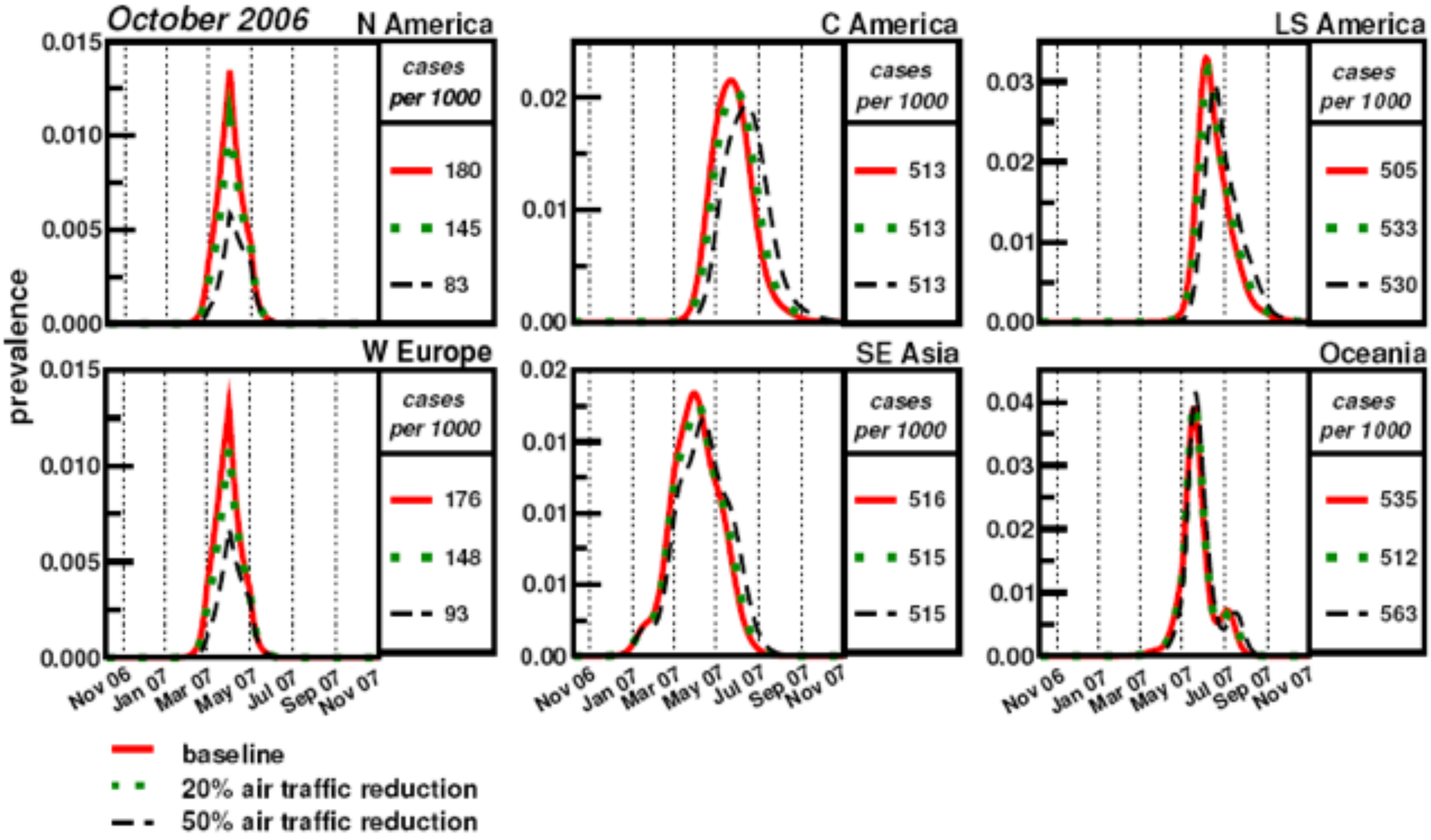
Prediction

months in advance

of the epidemic peak timing

see www.gleamviz.org

Travel limitations ?



Colizza, Barrat, Barthélemy, Valleron, Vespignani. *PLoS Medicine* (2007)

Analytical approach

=> Degree-based mean-field

Diffusion (random walk) between nodes

Reaction (SIS, SIR) inside each node

Degree-based mean-field approach: Diffusion

N nodes, W walkers

Node i => W_i walkers

$W = \Sigma_i W_i$

Degree block variables $$W_k = \frac{1}{N_k} \sum_{i|k_i=k} W_i$$

$$N_k = NP(k)$$ = number of nodes of degree *k*

Evolution equation:

$$\partial_t W_k(t) = -\underbrace{r_k W_k(t)}_{\text{Walkers going out of nodes}} + \underbrace{k \sum_{k'} \overbrace{d_{k'k}}^{\text{Diffusion rate along edges } k'\text{-}k} P(k'|k) W_{k'}(t)}_{\text{Walkers going into nodes}}$$

$r_k = k \sum_{k'} d_{kk'} P(k'|k)$

Degree-based mean-field approach: Diffusion

Simplest case: uniform diffusion $r_k = r; d_{k'k} = r/k'$

Uncorrelated random networks:

$$\partial_t W_k(t) = -rW_k(t) + \frac{k}{\langle k \rangle} \sum_{k'} P(k') r W_{k'}(t)$$

Stationarity => $\boxed{W_k(t) = \frac{k}{\langle k \rangle} \frac{W}{N}}$

Degree-based mean-field approach: Diffusion

Example of other diffusion rates $d_{kk'} = w_0 (kk')^\theta / T_k,\ r_k = r$

Uncorrelated random networks:

$$\partial_t W_k(t) = -r W_k(t) + r k^{1+\theta} \frac{w_0}{A\langle k \rangle} \sum_{k'} P(k') W_{k'}(t)$$

Stationarity =>

$$\boxed{W_k(t) = \frac{k^{1+\theta}}{\langle k^{1+\theta} \rangle} \frac{W}{N}}$$

Colizza, Vespignani, J. Theor. Biol. 251:450 (2008)

Degree-based mean-field approach: Diffusion

Diffusion rate keeping constant populations:
important in the perspective of modelling travel behaviours

Number of travellers between 2 subpopulations per unit time=fixed

Proba per unit time to go from *i* to *j*: $\frac{w_{ij}}{W_i}$

➡

$$\partial_t W_i = \sum_j W_j \frac{w_{ij}}{W_j} - W_i \sum_j \frac{w_{ij}}{W_i} = 0$$

Any population distribution is stationary

Degree-based mean-field approach: Diffusion

Diffusion rate keeping constant populations:
important in the perspective of modelling travel behaviours

In the degree-based framework

$$d_{kk'} = \frac{w_{kk'}}{W_k} \quad (w_{kk'} = w_{k'k}) \qquad\qquad r_k = k \sum_{k'} d_{kk'} P(k'|k)$$

$$\begin{aligned}
\partial_t W_k(t) &= -r_k W_k(t) + k \sum_{k'} d_{k'k} P(k'|k) W_{k'}(t) \\
&= -k \sum_{k'} \frac{w_{kk'}}{W_k} P(k'|k) W_k + k \sum_{k'} w_{k'k} P(k'|k) \\
&= 0
\end{aligned}$$

Any population distribution is stationary

Colizza, Vespignani, J. Theor. Biol. 251:450 (2008)

Degree-based mean-field approach: SIS

In each node i: S_i susceptible, I_i infectious, $W_i = S_i + I_i$

Degree block variables $$S_k = \frac{1}{N_k} \sum_{i|k_i=k} S_i \qquad I_k = \frac{1}{N_k} \sum_{i|k_i=k} I_i$$

Each time step: 2 processes

1- reaction

2- diffusion

Colizza, Vespignani, J. Theor. Biol. 251:450 (2008)

Degree-based mean-field approach: SIS

Each time step: 2 processes

$$\Gamma_k = S_k I_k / W_k$$

1- reaction $I_k \to I_k - \mu I_k + \beta \Gamma_k$

2- diffusion
$$\begin{aligned} I_k &\to (I_k - \mu I_k + \beta\Gamma_k)(1 - r_k) \\ &+ k \sum_{k'} P(k'|k) d_{k'k}((1-\mu) I_{k'} + \beta \Gamma_{k'}) \end{aligned}$$

Uniform diffusion or diffusion with constant populations
=> epidemic threshold $\beta/\mu = 1$

Colizza, Vespignani, J. Theor. Biol. 251:450 (2008)

Degree-based mean-field approach: SIR case

$$\beta/\mu > 1$$

- Zero diffusion: epidemics confined in first subpopulation
- Infinite diffusion: population well-mixed

⇓

expect a transition between

- confined epidemics at low diffusion rates
- global invasion at large diffusion rates

NB: for SIS, as soon as non-zero diffusion, global invasion as there is a stationary state

Colizza, Vespignani, J. Theor. Biol. 251:450 (2008)

Degree-based mean-field approach: SIR case

Problem:

Can be very small: travel of *fractions* of individuals

$$\begin{aligned}\partial_t I_k &= -r_k I_k + (-\mu I_k + \beta\Gamma_k)(1 - r_k) \\ &+ k\sum_{k'} P(k'|k)\, d_{k'k}\big((1-\mu)I_{k'} + \beta\Gamma_{k'}\big)\end{aligned}$$

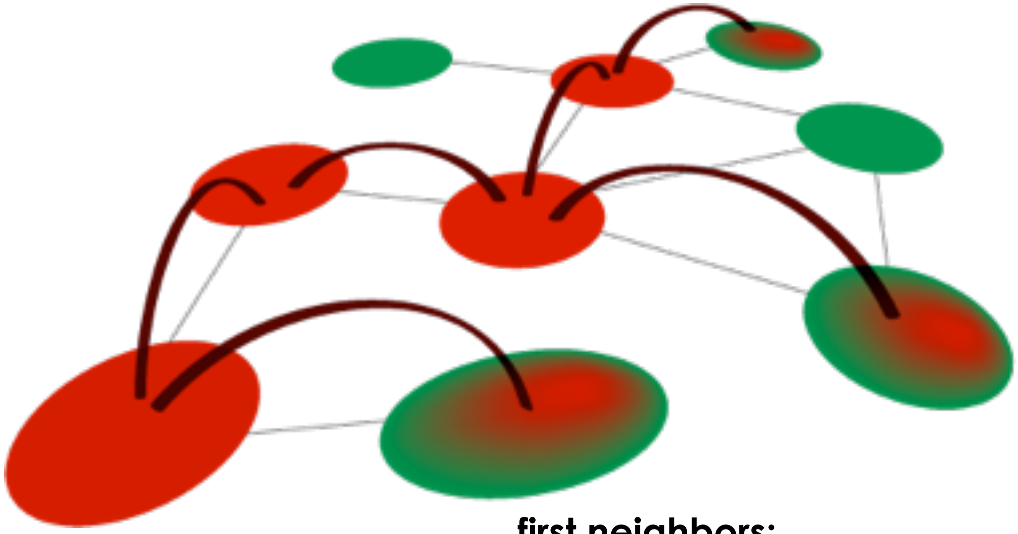
continuous approximation cannot capture the global invasion threshold

need to take into account **discreteness** & **stochasticity**

Colizza, Vespignani, J. Theor. Biol. 251:450 (2008)

Invasion: branching process

$D_k^0, D_k^1, \ldots$ # of ***diseased*** nodes (i.e., with at least one infected individual) of degree *k*, at generation *n=0, 1, ...*



seed:
generation n=0

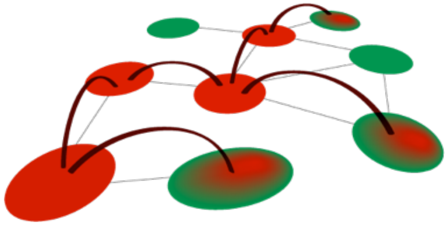
first neighbors:
generation n=1

Colizza, Vespignani, J. Theor. Biol. 251:450 (2008)

Invasion: branching process

$$D_k^n = \sum_{k'} D_{k'}^{n-1} \overbrace{(k'-1)}^{\text{possible paths of infection}} \overbrace{P(k|k')}^{\text{proba connection k'-k}}$$

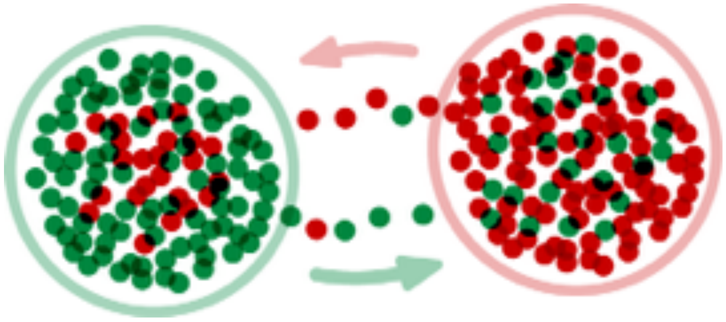
$$\times \underbrace{\left(1 - \frac{D_k^{n-1}}{N_k}\right)}_{\text{probability of finding uninfected subpopulation}} \underbrace{\left[1 - \left(\frac{1}{R_0}\right)^{\lambda_{k'k}}\right]}_{\text{proba of outbreak in pop. k= 1 - proba of extinction for a seed of } \lambda_{k'k} \text{ individuals}}$$



$\lambda_{k'k}$ Number of infectious individuals that move into population *k* during the outbreak in population of degree *k'*

Colizza, Vespignani, J. Theor. Biol. 251:450 (2008)

from micro to macro scale



\# of new seeds from *k'* to *k*
(time-scale separation):

$$\lambda_{k'k} = \underbrace{d_{k'k}}_{\text{diffusion rate}} \frac{\overbrace{\alpha N_{k'}}^{\textit{total \# of infected generated}}}{\underbrace{\mu}_{\text{rate } I \rightarrow R}}$$

Ex of diffusion rates:

$$d_{k'k} = p\, w_0\, (k'k)^{\theta} / T_{k'}$$

$$d_{k'k} = w_0\, (kk')^{\theta} / \mathbf{N}_{k'}$$

(stationary populations ind. of diffusion process))

Colizza, Vespignani, J. Theor. Biol. 251:450 (2008)

For R_0 close to 1, uncorrelated networks and at short times:

$$D_k^n = (R_0 - 1)\frac{k^{1+\theta}P(k)}{\langle k \rangle}\frac{w_0\alpha}{\mu}\underbrace{\sum_{k'} D_{k'}^{n-1} k'^{\theta}(k'-1)}_{\Theta^{n-1}}$$

$$\Downarrow$$

$$\Theta^n = \underbrace{(R_0 - 1)\frac{\langle k^{2+2\theta} \rangle - \langle k^{1+2\theta} \rangle}{\langle k \rangle}\frac{w_0\alpha}{\mu}}_{> 1 \Leftrightarrow \text{ global invasion}}\Theta^{n-1}$$

Colizza, Vespignani, J. Theor. Biol. 251:450 (2008)

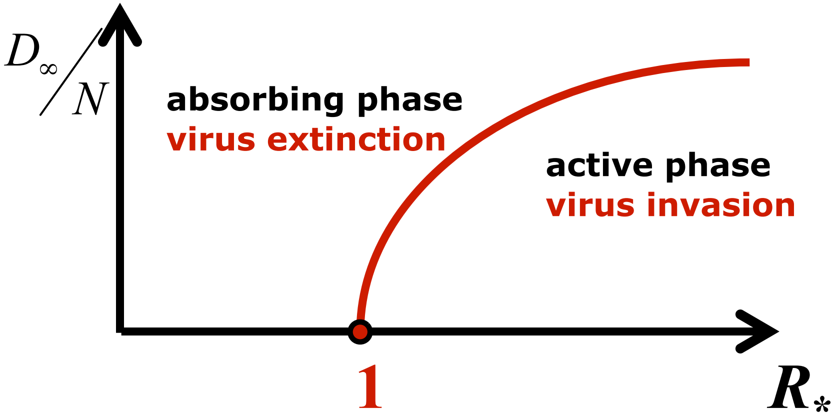
Global invasion threshold

$$R_* = (R_0 - 1)\frac{\langle k^{2+2\theta}\rangle - \langle k^{1+2\theta}\rangle}{\langle k\rangle}\frac{w_0\alpha}{\mu}$$

Ex: SIR, $\alpha \sim 2(R_0-1)/R_0^2$

$$w_{0c} = \frac{\mu R_0^2}{2(R_0-1)^2}\frac{\langle k\rangle}{\langle k^{2+2\theta}\rangle - \langle k^{1+2\theta}\rangle}$$

phase transition in mobility



Real-world network: w_0 10 times larger than w_{0c} !!!

➡ Explains empirical results!!

Colizza & Vespignani, PRL (2007), JTB (2008)

Going beyond

- Population structure (age/gender) and travel behaviours

(Apolloni et al., BMC ID 2013)

- Length of stay at destination

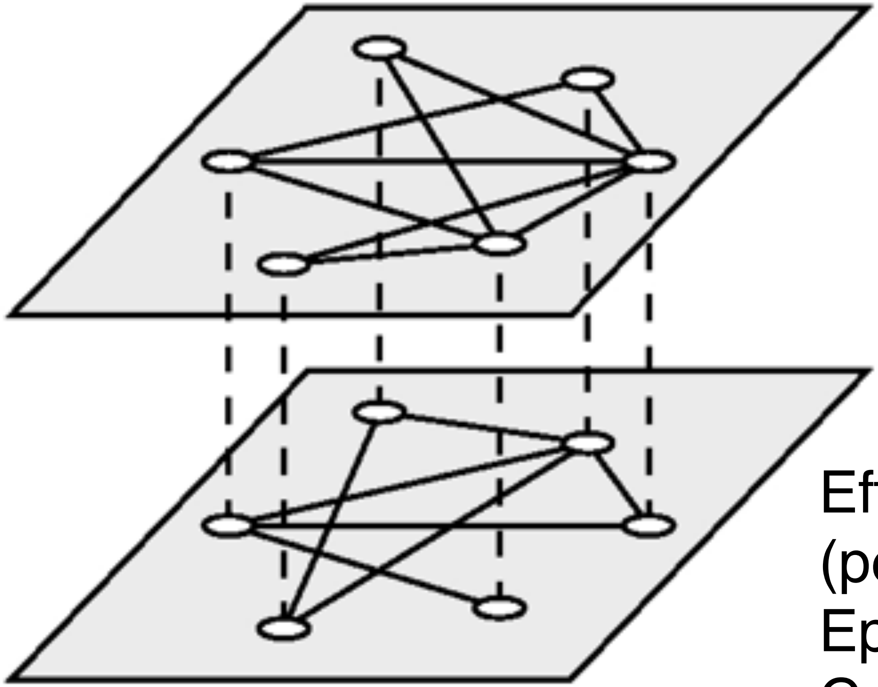
(Poletto et al., J. Th. Biol. 2013)

- Change of behaviour

(Meloni et al., Sci. Rep. 2011)

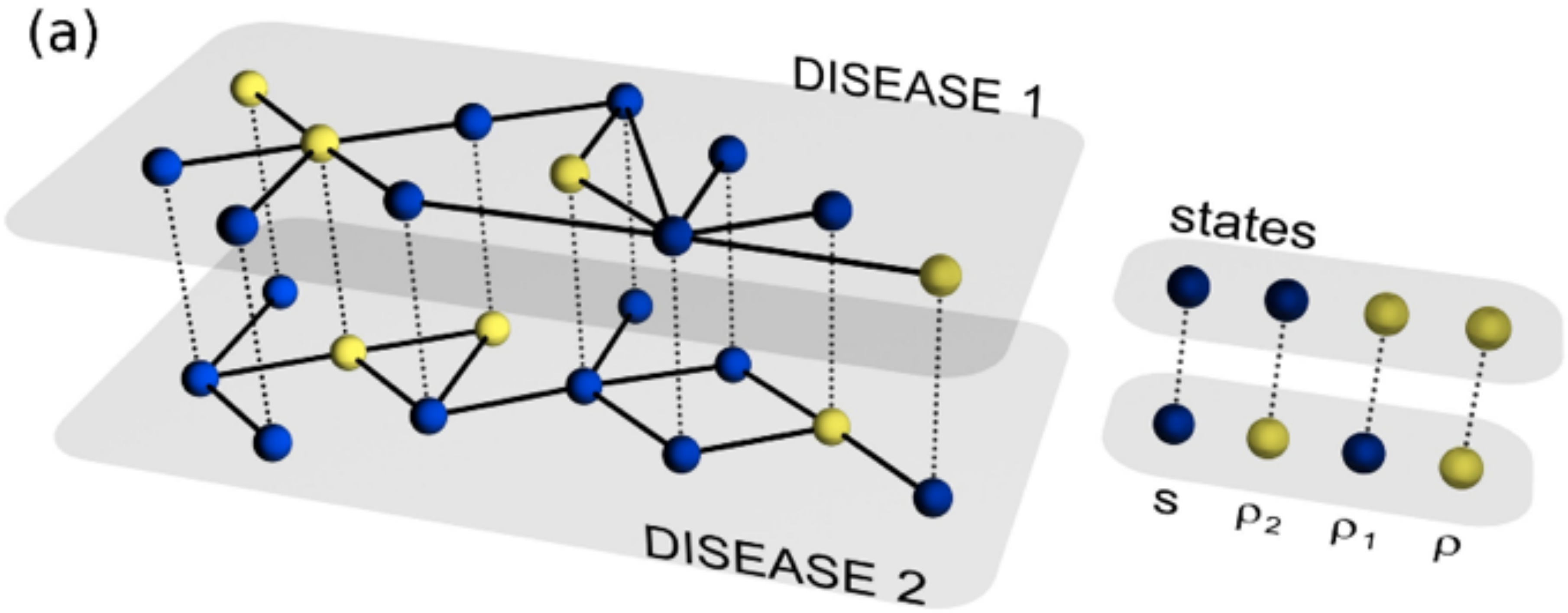
Epidemics in multiplex networks

Interdependent networks (power-grid - communication/computer network)
Layers of social networks
Different transportation networks
…



Effect of coupling on cascading failures (percolation processes)
Epidemics on multiplex networks
Cooperation in multiplex social networks
…

Epidemics in multiplex networks



Modification of infectiousness of disease (a)
if spreader or susceptible are infected with disease (b

Modification of recovery rate for disease (a) if
infectious is also infected with disease (b)

→

mutual enhancement
or
partial cross-immunity

Sanz et al., arXiv:1402.4523

Epidemics in time-varying networks

Networks= (often) dynamical entities

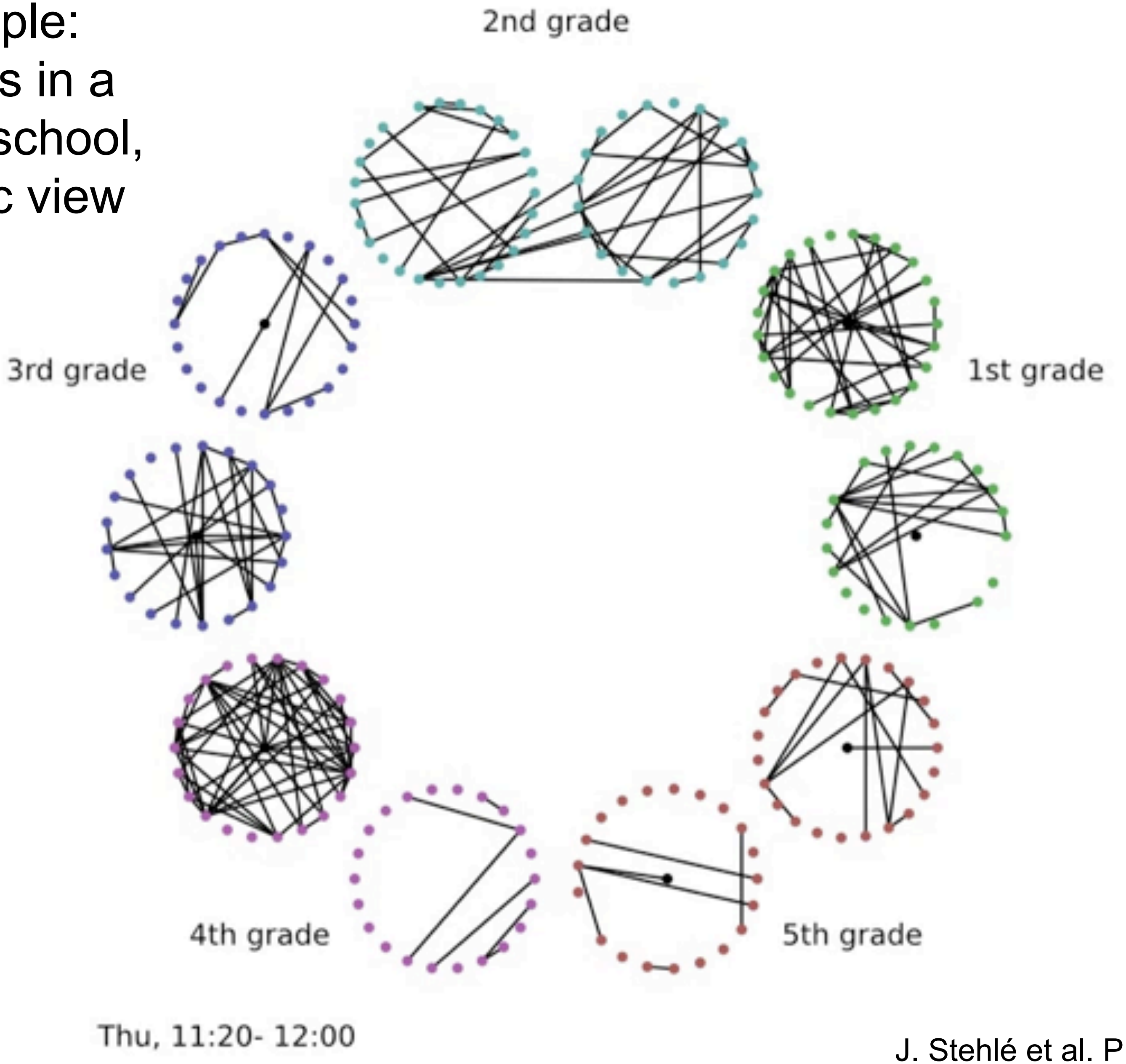
(communication, social networks, online networks, transport networks, etc…)

- Which dynamics?
- Characterization?
- Modeling?
- Consequences on dynamical phenomena? (e.g. epidemics, information propagation…)

Time-varying networks: often represented by aggregated views

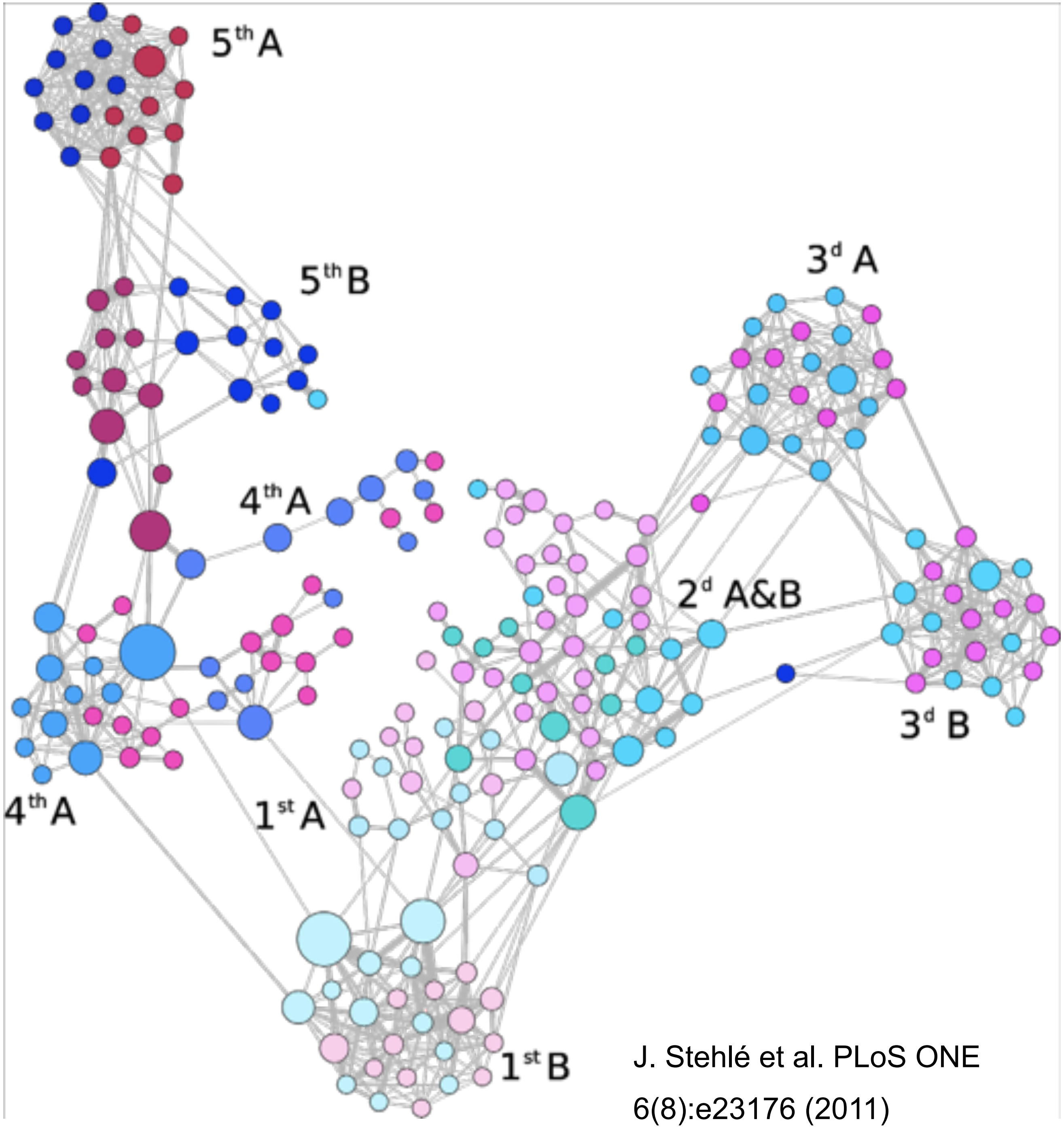
- Lack of data
- Convenience

Example: contacts in a primary school, dynamic view



J. Stehlé et al. PLoS ONE 6(8):e23176 (2011)

Example: contacts in a primary school, aggregated view



J. Stehlé et al. PLoS ONE 6(8):e23176 (2011)

Definition: temporal network

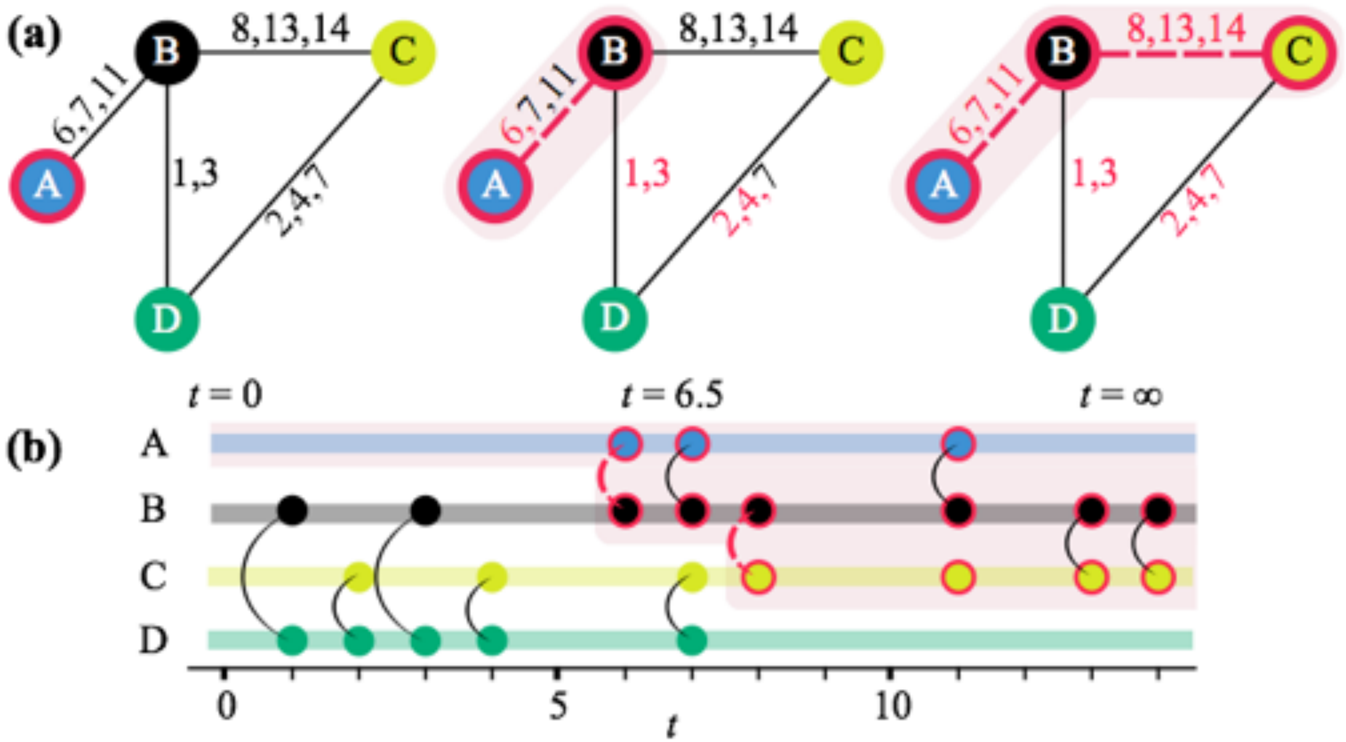
Temporal network: T=(V,S)

- V=set of nodes
- S=set of event sequences assigned to pairs of nodes

$$s_{ij} \in S : \, s_{ij} = \{(t_{ij}^{s,1}, t_{ij}^{e,1}) \cdots (t_{ij}^{s,\ell}, t_{ij}^{e,\ell})\}$$

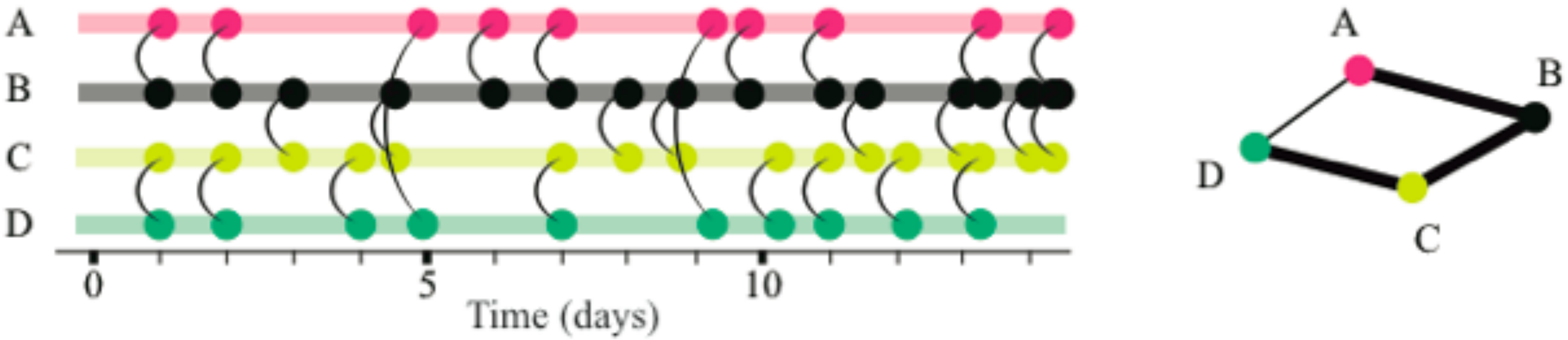
Other representation:
time-dependent adjacency matrix:
a(i,j,t)= 1 <=> i and j connected at time t

Reachability in temporal networks



Review Holme-Saramaki, Phys. Rep. (2012), arXiv:1108.1780

Aggregation of temporal network



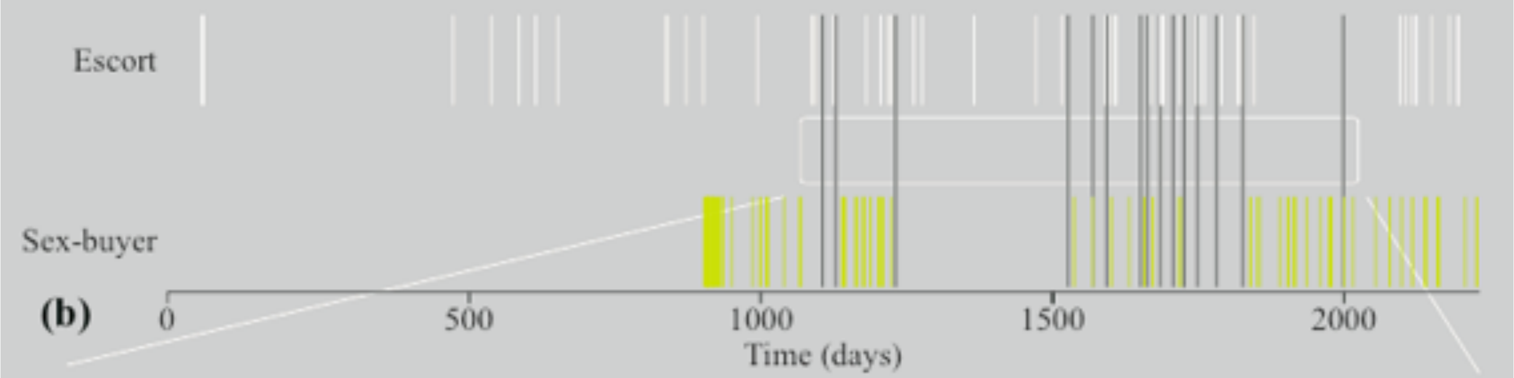
$$w_{ij} = \int_{t_{min}}^{t_{max}} a(i,j,t)dt$$

NB: enough information if underlying process is Poissonian

Review Holme-Saramaki, Phys. Rep. (2012), arXiv:1108.1780

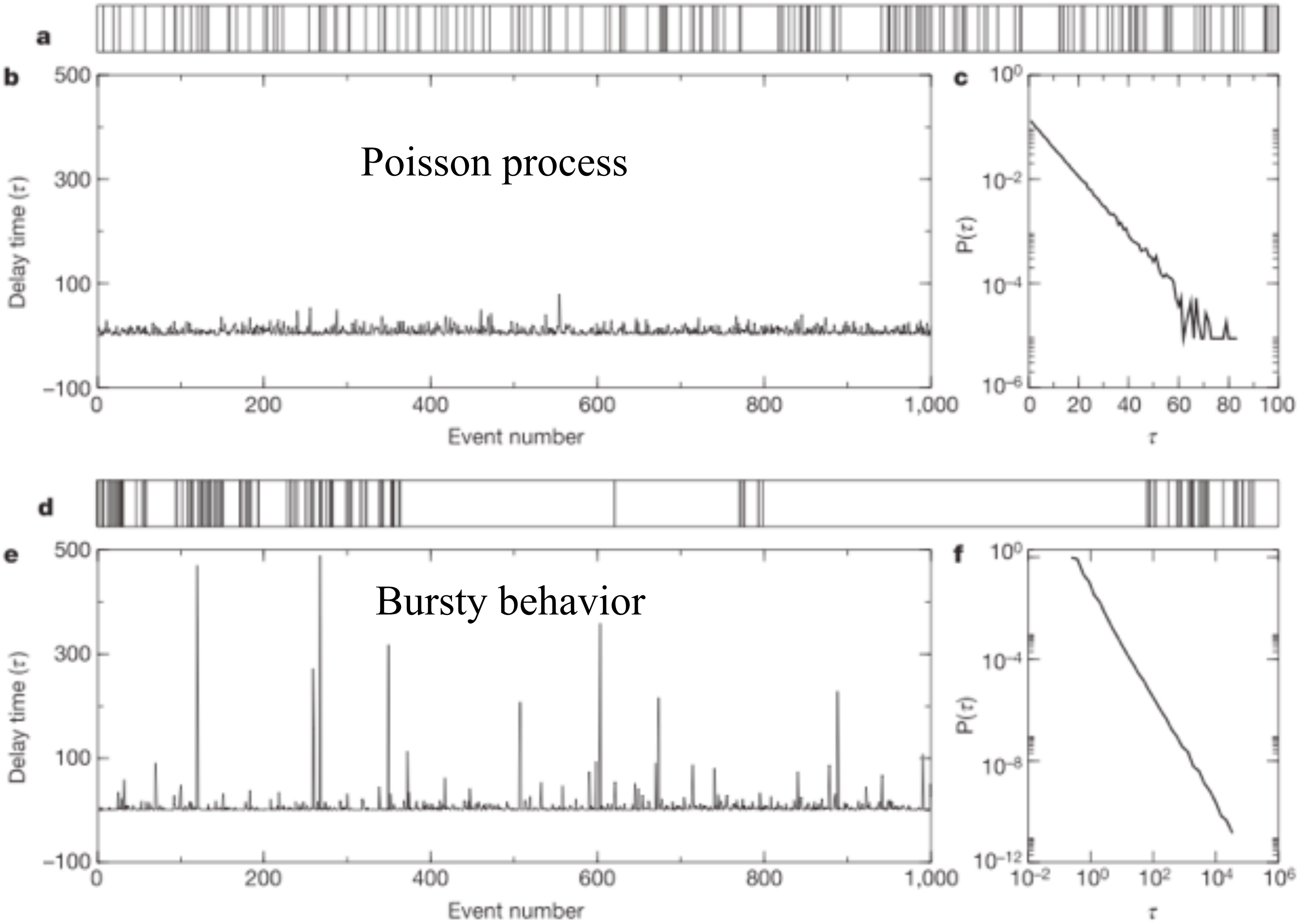
Aggregation of temporal network

Temporal behavior most often non-Poissonian
=> aggregate view hides important temporal patterns



Review Holme-Saramaki, Phys. Rep. (2012), arXiv:1108.1780

Burstiness



Poisson process

Bursty behavior

A.-L. Barabasi, Nature (2006)

Generalization of definitions to temporal networks

Reachability issue
=> **time respecting path** ("journey")
    => set of influence of a node
    => temporal connectivity (similar to case of directed graphs)

Path length => concept of **shortest paths**
Time respecting path duration => concept of **fastest journey**

Temporal motifs

Centrality measures

(…)

Review Holme-Saramaki, Phys. Rep. (2012), arXiv:1108.1780

Complex temporal characteristics

- burstiness
  - non-Poissonian inter-event distributions
  - power-law temporal correlations
- heterogeneity of event durations
  - single events
  - aggregated durations (weights in aggregated networks)
- stationarity of statistical features
- daily, weekly, and organizational rhythms
- weight-topology correlations
- topology-activity correlations (e.g., school)

…

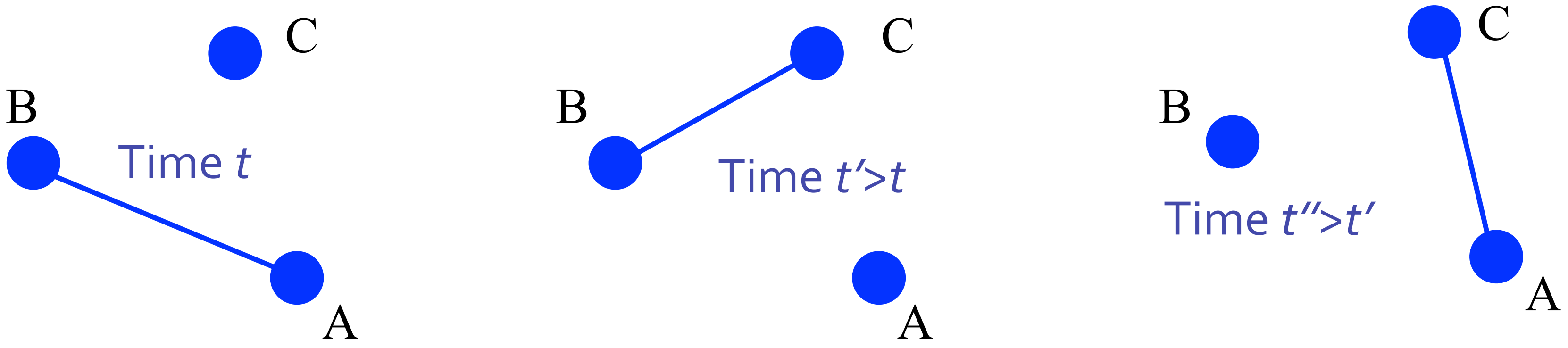
Temporal networks

- Generalization of concepts?
- Centrality of a node?
- Temporal communities?
- Models for temporal networks?
- Impact of temporal features on dynamical processes?

Toy spreading processes on dynamical networks

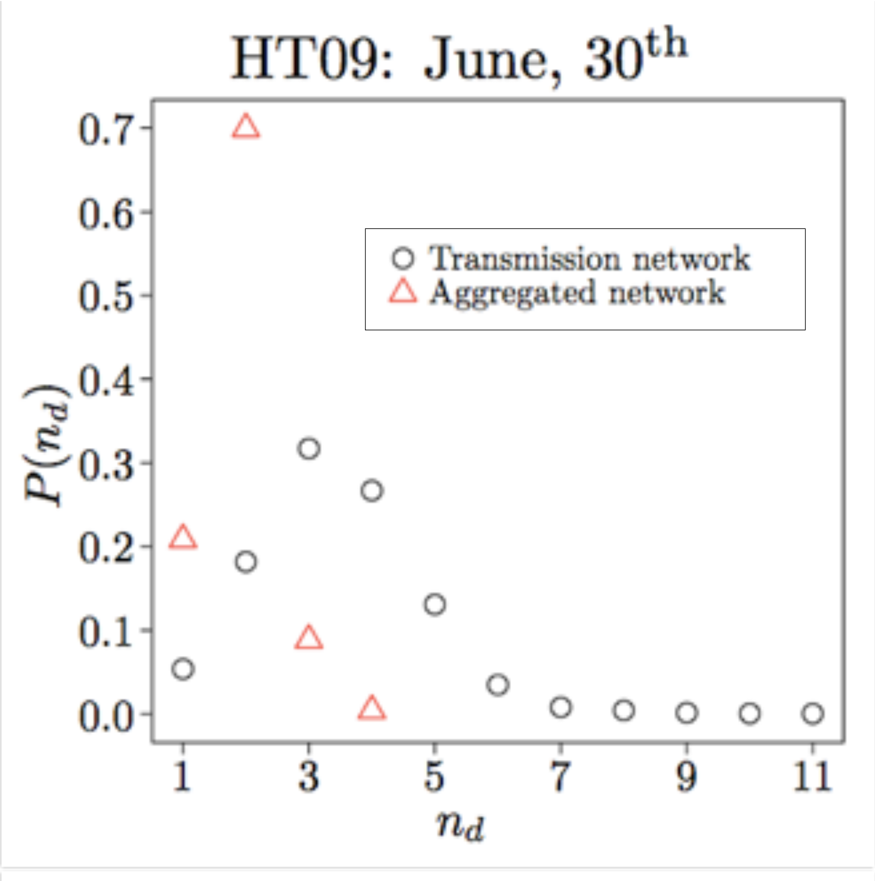
- deterministic SI process

- **fastest paths ≠ shortest paths**

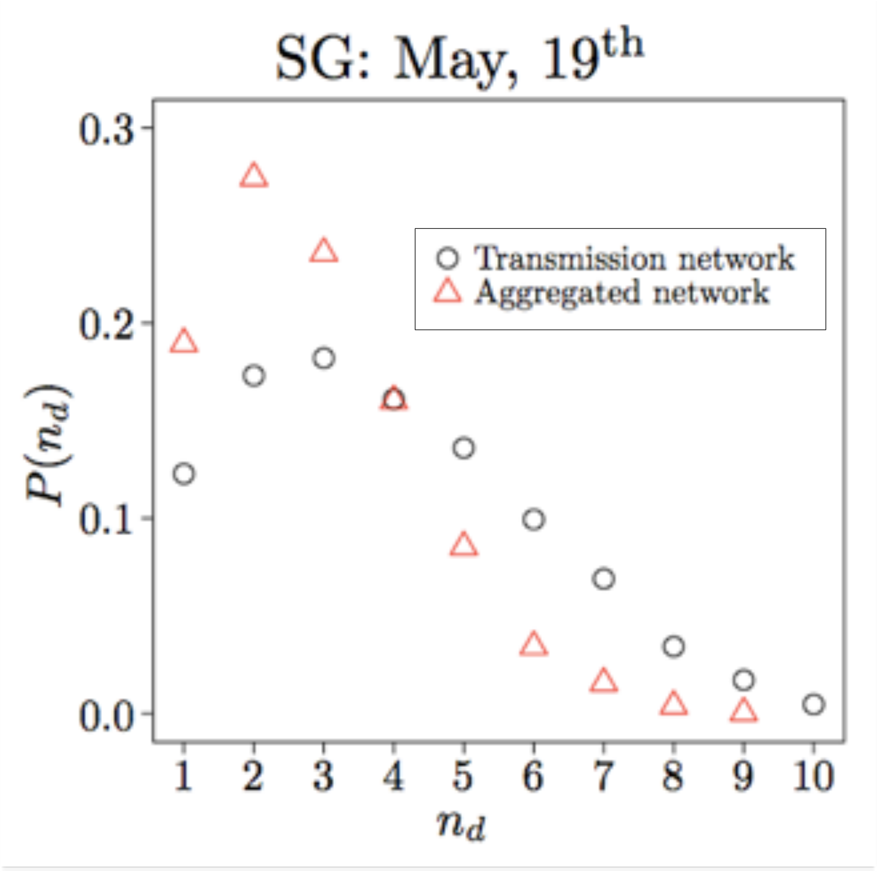


Fastest path= A->B->C
Shortest path= A-C

Example: shortest vs fastest paths in a temporal contact network

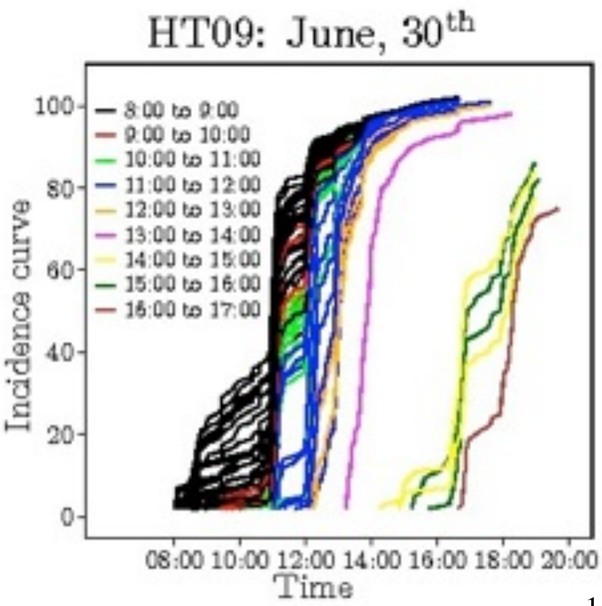


Conference

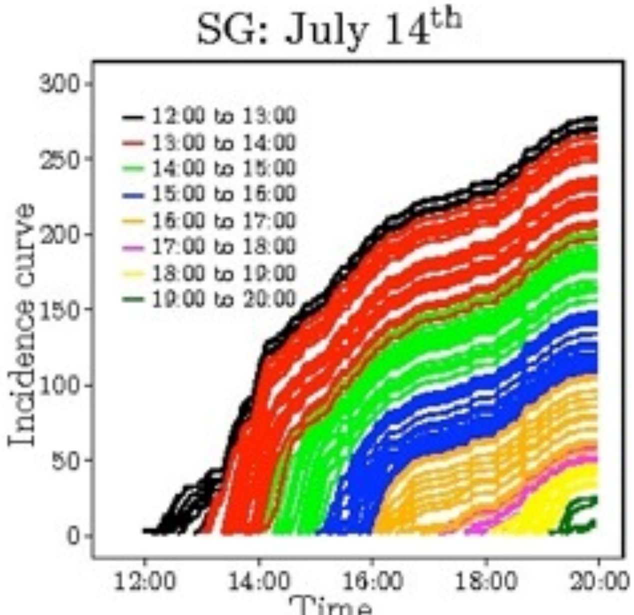


Museum

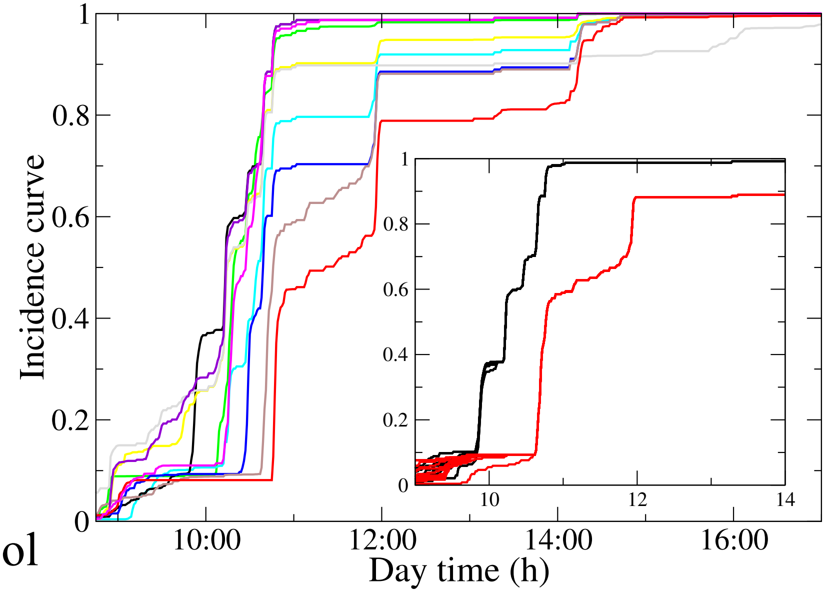
Example: deterministic SI in temporal contact networks



Conference

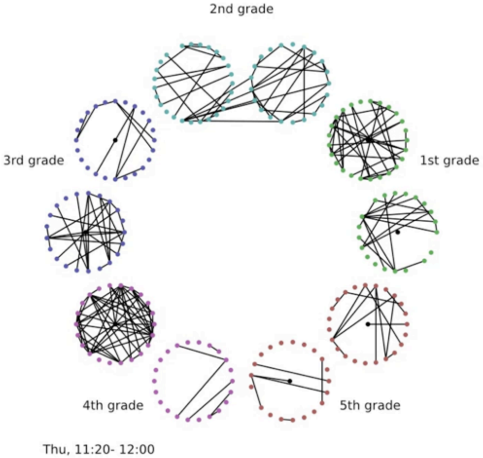


Museum



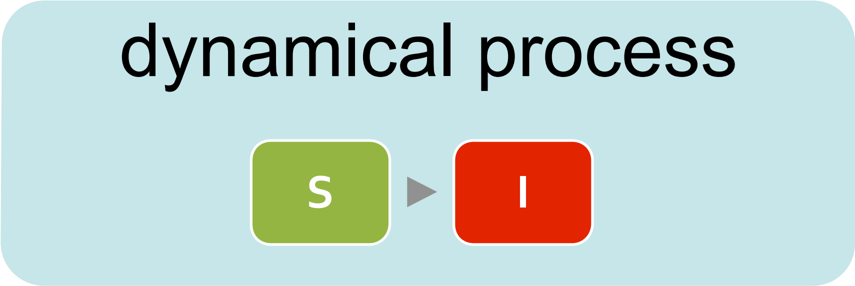
School

>(Toy) spreading processes on dynamical networks



Thu, 11:20- 12:00

+



dynamical process

S ▶ I

Use of null models to reveal the role of the temporal aspects

Mobile phone data:
- community structure (C)
- weight-topology correlations (W)
- burstiness on single links (B)
- daily patterns (D)
- event-event correlations between links (E)

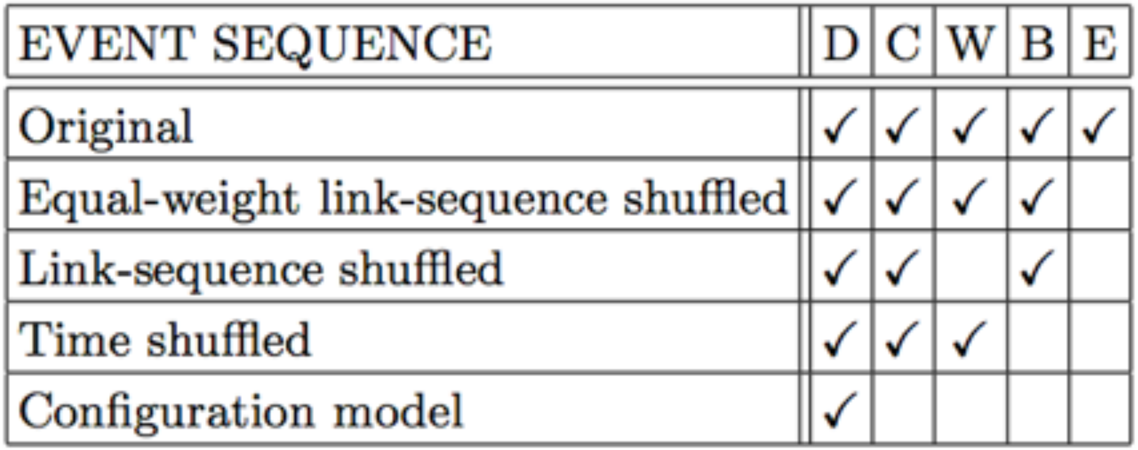
Effects of the different ingredients?

Use series of **null models**!

M. Karsai et al., Small But Slow World: How Network Topology and Burstiness Slow Down Spreading, Phys. Rev. E (2011).

- community structure (C)
- weight-topology correlations (W)
- burstiness on single links (B)
- daily patterns (D)
- event-event correlations between links (E)

Null models

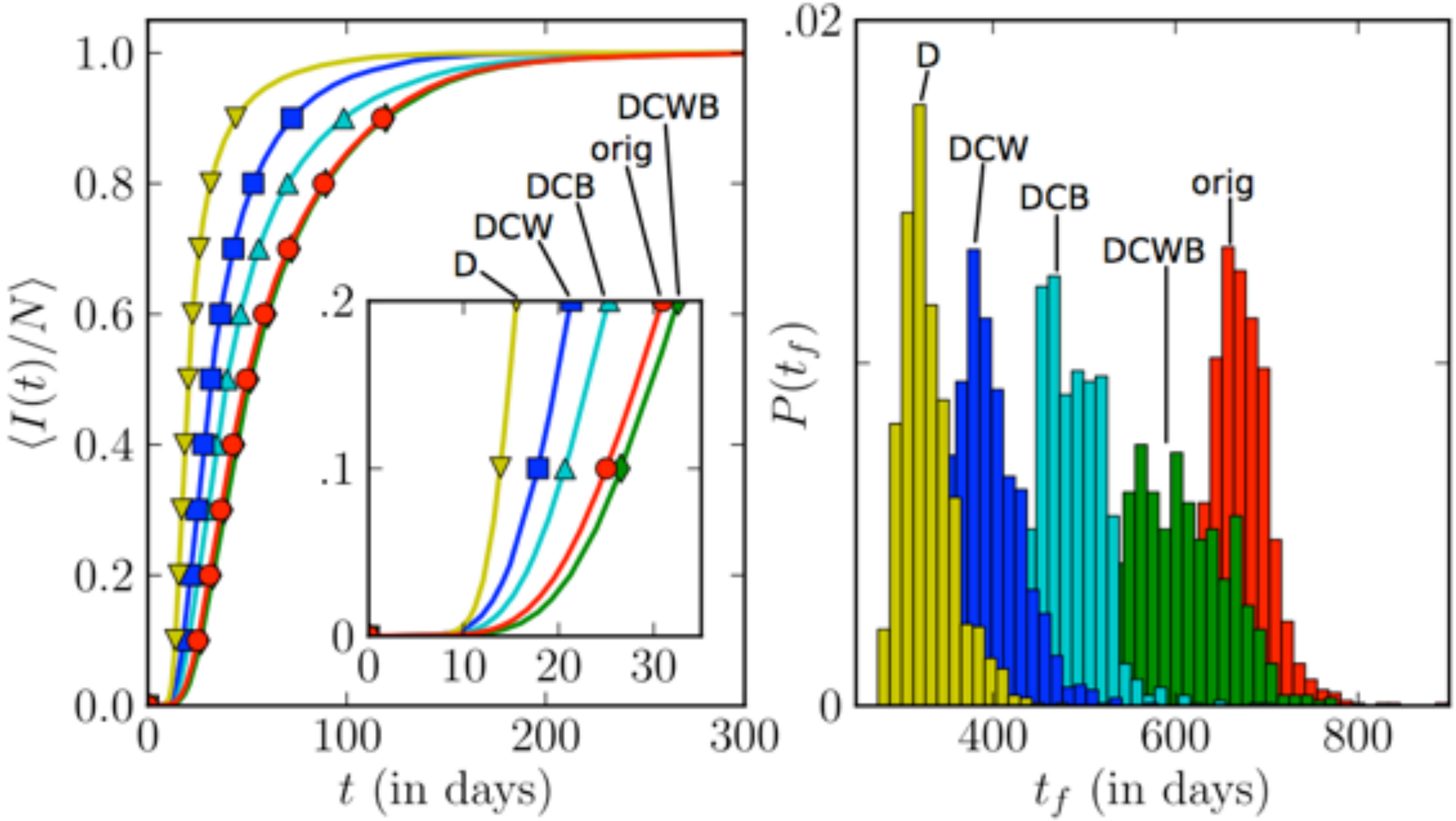


EVENT SEQUENCE	D	C	W	B	E
Original	✓	✓	✓	✓	✓
Equal-weight link-sequence shuffled	✓	✓	✓	✓	
Link-sequence shuffled	✓	✓		✓	
Time shuffled	✓	✓	✓		
Configuration model	✓				

M. Karsai et al., Small But Slow World: How Network Topology and Burstiness Slow Down Spreading, Phys. Rev. E (2011).

Mobile phone data

- community structure (C)
- weight-topology correlations (W)
- burstiness on single links (B)
- daily patterns (D)
- event-event correlations between links (E)



Bursty dynamics slows down spreading

M. Karsai et al., Small But Slow World: How Network Topology and Burstiness Slow Down Spreading, Phys. Rev. E (2011).
Kivela et al, Multiscale Analysis of Spreading in a Large Communication Network, arXiv:1112.4312

More results

Rocha et al., PLOS Comp Biol (2011)
- data: temporal network of sexual contacts
- temporal correlations accelerate outbreaks

Pan & Saramaki, PRE (2011)
- data: mobile phone call network
- slower spread when correlations removed

Miritello et al., PRE (2011)
- data: mobile phone call network
- burstiness decreases transmissibility

Takaguchi et al., PLOS ONE (2013)
- data: contacts in a conference; email
- threshold-based spreading model
- burstiness accelerate spreading

Rocha & Blondel, PLOS Comp Biol (2013)
- model with tuneable distribution of inter-event times
(no correlations)
- burstiness => initial speedup, long time slowing down

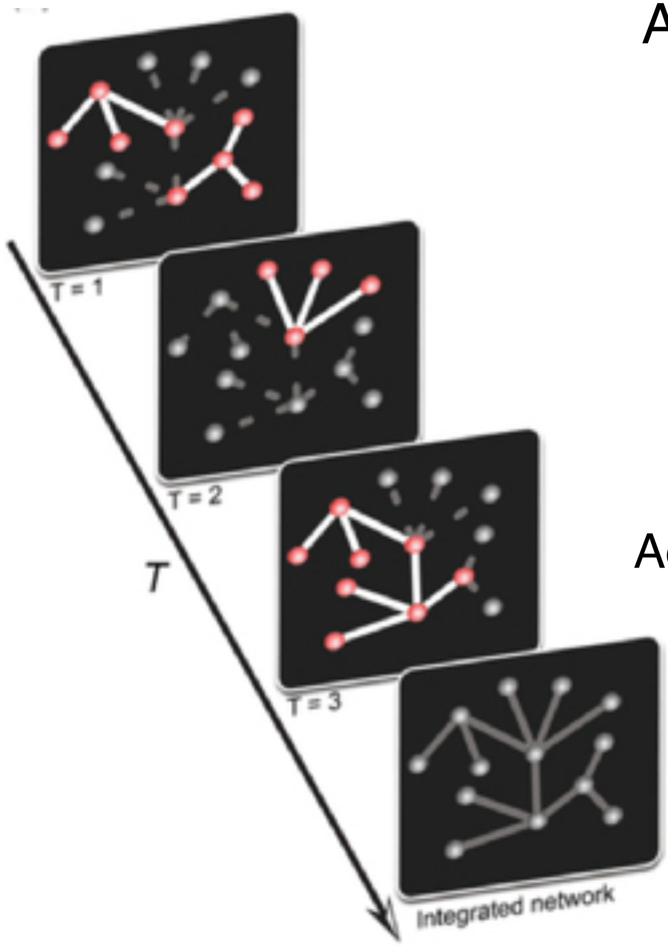
Still somewhat unclear picture

Results

- depend on data set
- depend on spreading model
- generally
    - burstiness slows down spreading
    - correlations (e.g., temporal motifs) favors spreading
    - role of turnover
    - +: effect of static patterns

SIS model on activity-driven network

Model: N nodes, each with an “activity” a, taken from a distribution F(a)



At each time step:

- node i active with probability a(i)
- each active node generate m links to other randomly chosen nodes
- iterate with no memory

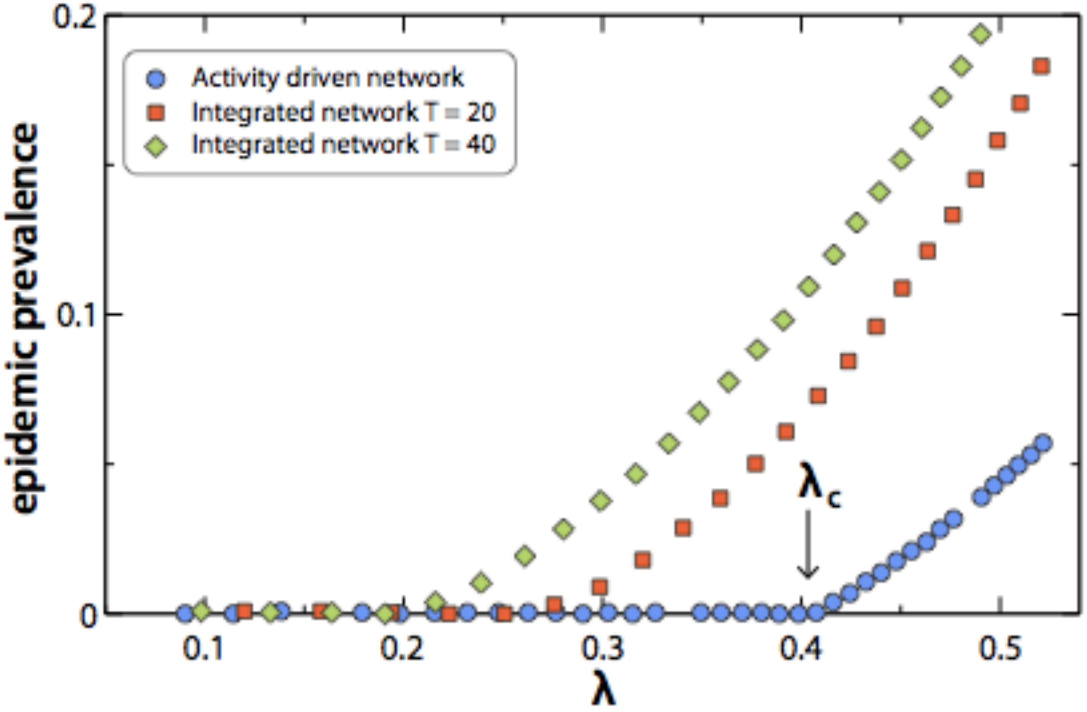
Activity-based mean-field theory:

$$I_a^{t+\Delta t} = -\mu \Delta t I_a^t + I_a^t + \lambda m \left(N_a^t - I_a^t\right) a \Delta t \int da' \frac{I_{a'}^t}{N} + \lambda m \left(N_a^t - I_a^t\right) \int da' \frac{I_{a'}^t a' \Delta t}{N},$$

Perra et al., Sci. Rep. (2012)

SIS model on activity-driven network

Epidemic threshold: $$\lambda_c = \frac{1}{m(\langle a \rangle + \sqrt{\langle a^2 \rangle})}$$



Perra et al., Sci. Rep. (2012)

Immunization strategies

=> take into account temporal structure

Lee et al., PLOS ONE (2012)
=>inspired by “acquaintance protocol” in static networks)

- “*Recent*”: choose a node at random, immunize its most recent contact
- “*Weight*”: choose a node at random, immunize its most frequent contact in a previous time-window

Starnini et al., JTB (2012)

- aggregate network on [0,T]
- compare strategies
    - immunize nodes with highest k or BC in [0,T]
    - immunize random acquaintance (on [0,T])
    - recent, weight strategies
- vary T
- find saturation of efficiency as T increases

Liu et al., arXiv:1309:7031 (activity-driven network model=>analytics)

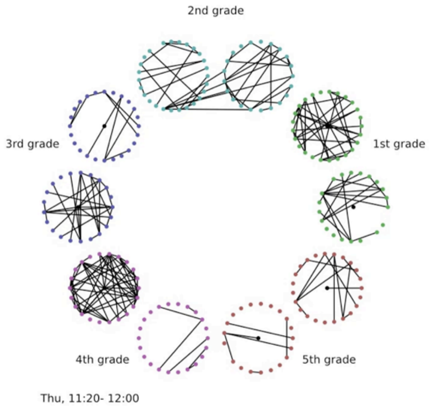
- target nodes with largest activity
- random neighbour (over an observation time T) of random node

>Representing data for data-driven numerical simulations

- data (e.g. contacts) measured in one specific context at one specific time

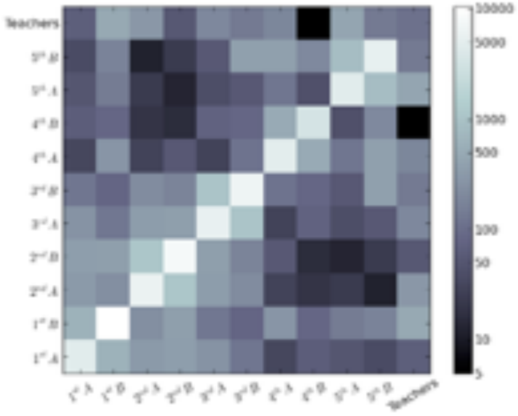
- need to perform numerical simulations that can give information on a potential spread at a different time (in similar context)

How much detail to inform the models?



Detailed dynamic network

- very detailed ✔
- very realistic ✔
- takes into account individual heterogeneities of behavior ✔
- very specific (context+period), not easy to generalize ✗

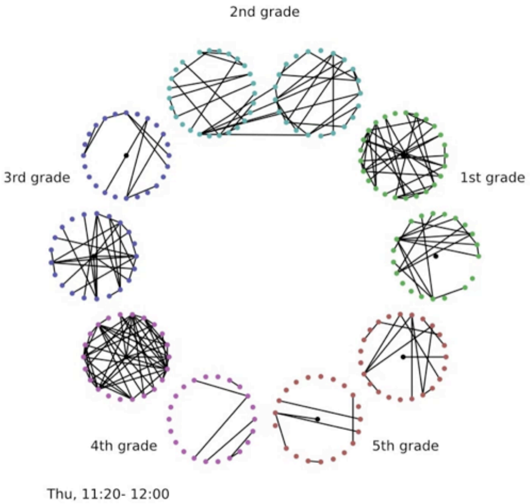


Contact matrix

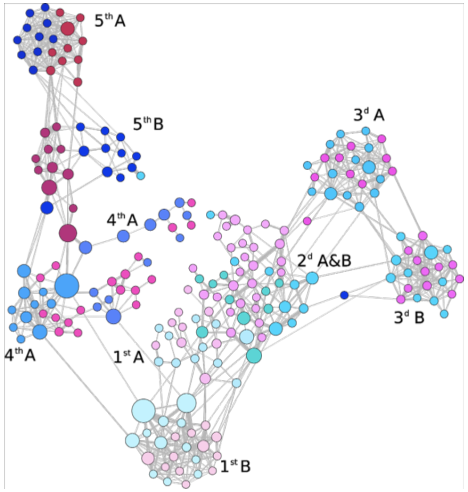
- coarse-grained ✗
- fully connected structure ✗
- only heterogeneities between groups ✗
- very easy to generalize ✔

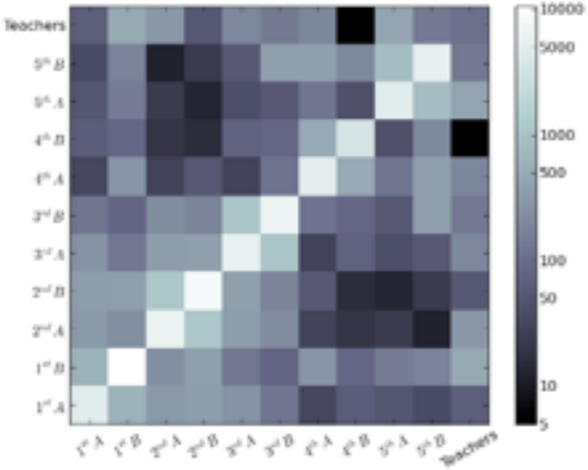
“synopsis” of dynamic network data

Temporal network



Static network

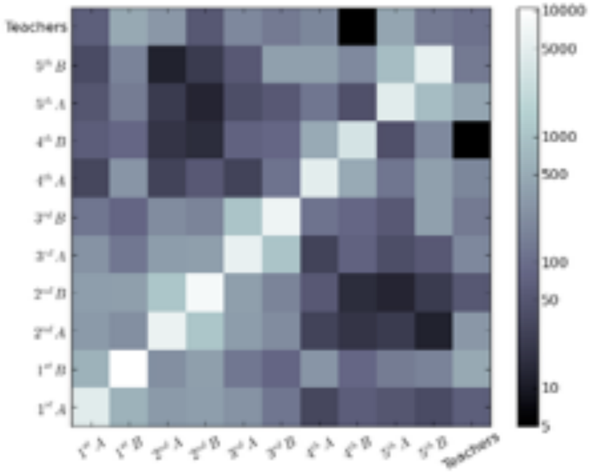




Contact matrix
(underlying fully connected assumption + no within-class heterogeneity)

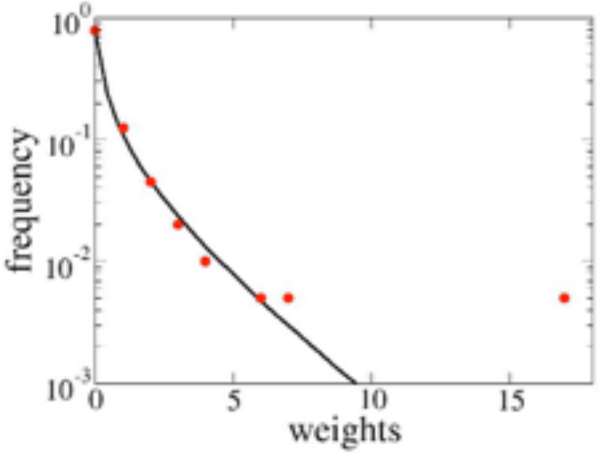
“synopsis” of dynamic network data

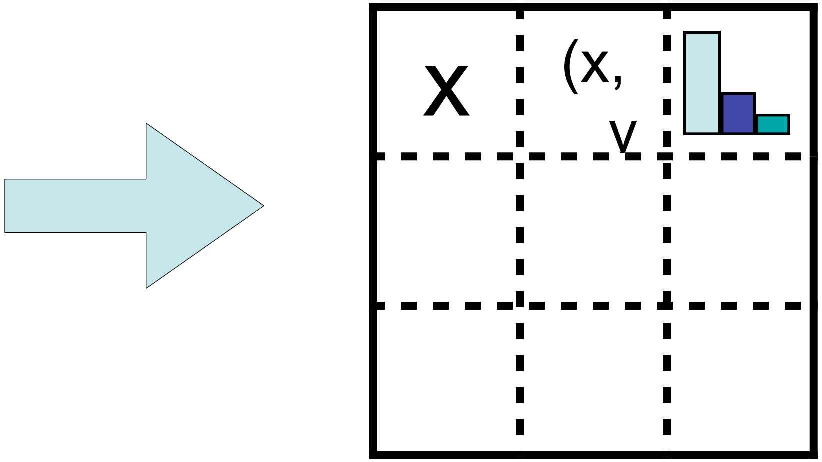
Role-based structure



+

Heterogeneities





Contact matrix of distributions:

-role based

-takes heterogeneities into account

Evaluating representations

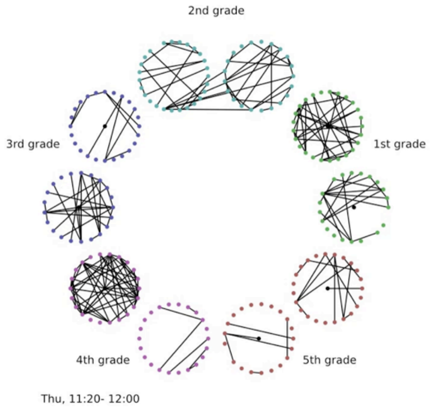
data representation

+

epidemic model

S ▶ E ▶ I ▶ R

vs



+

epidemic model

S ▶ E ▶ I ▶ R

>A concrete example: contact patterns in a hospital

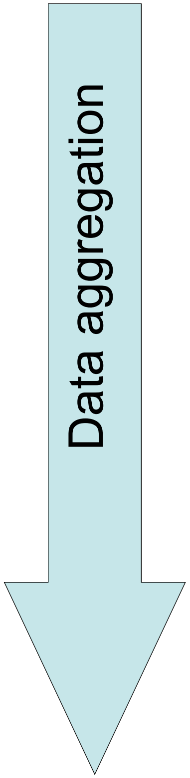
Network representations

Construction of 3 networks:

1. Dynamic network (DYN):
   Real sequence of successive contacts

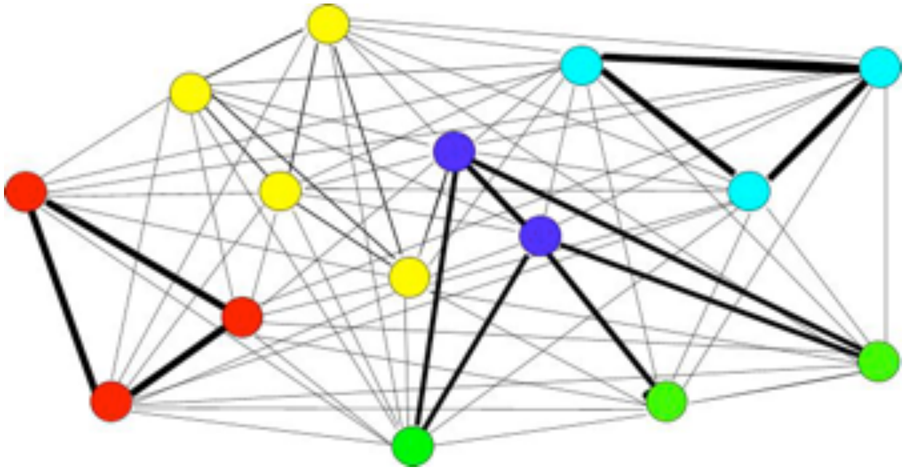
2. Heterogeneous network (HET):
   1-day aggregated network
   A—B if A and B have been in contact
   W_{AB} = cumulative duration of the contacts A-B

3. Homogeneous network (HOM):
   1-day aggregated network
   A—B if A and B have been in contact
   W_{AB} = average cumulative duration

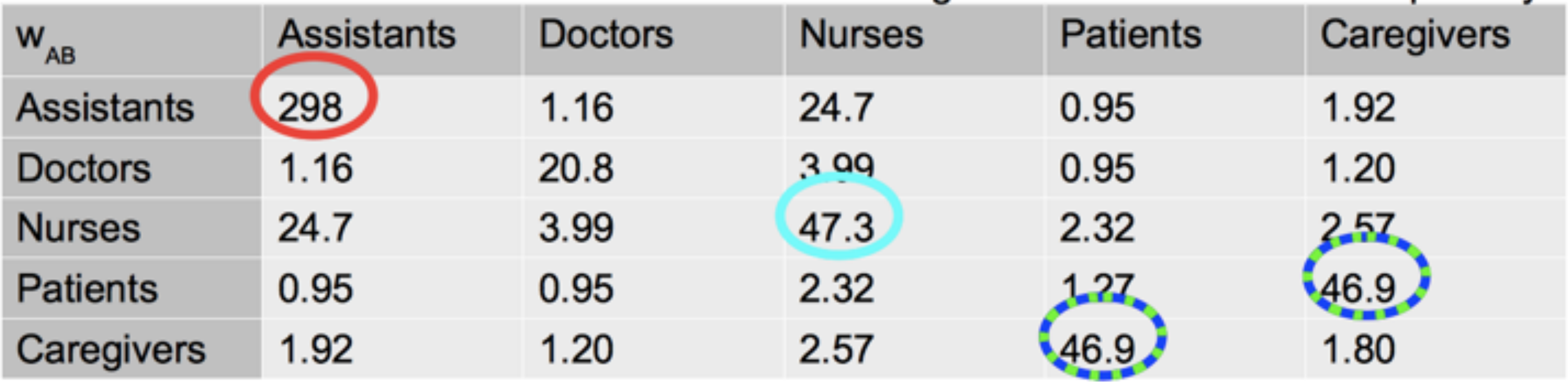


Networks:
- Take into account network structure at the individual level
- Difficult to generalize

4. Contact matrix



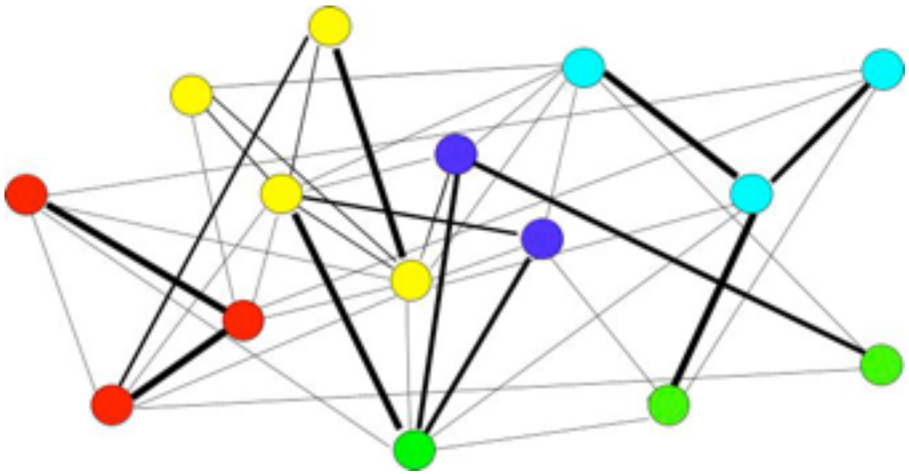
Average contact time in seconds per day



w_{AB}	Assistants	Doctors	Nurses	Patients	Caregivers
Assistants	298	1.16	24.7	0.95	1.92
Doctors	1.16	20.8	3.99	0.95	1.20
Nurses	24.7	3.99	47.3	2.32	2.57
Patients	0.95	0.95	2.32	1.27	46.9
Caregivers	1.92	1.20	2.57	46.9	1.80

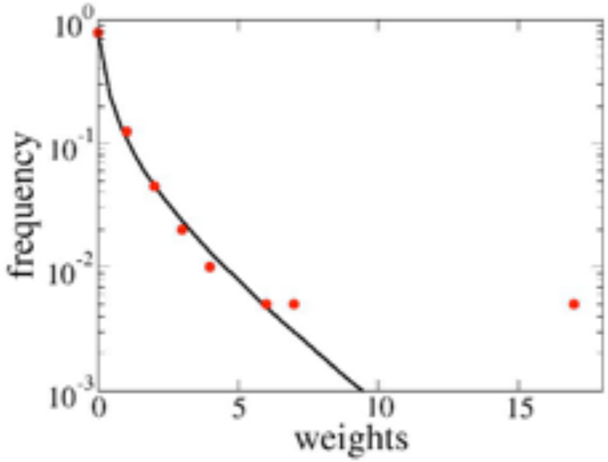
- Underlying fully connected network structure
- Takes into account role structure
- Average temporal information, no heterogeneities within each role
- Easy to generalize

5. Novel representation: Contact matrix of distributions



a. Fit each role-pair distribution of weights
(using negative binomials)
b. Create a network in which weights are drawn
from the fitted distribution (NB: including zero weights)

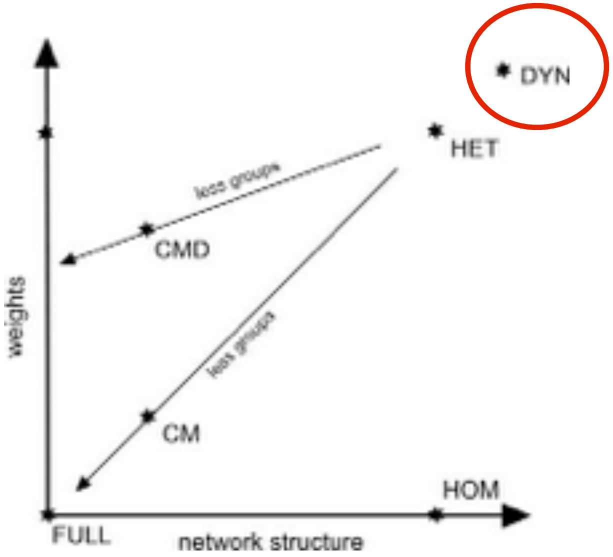
Example: Assistant-Doctor



- Underlying realistic network structure
- Takes into account role structure
- Takes into account heterogeneities within each role
- Easy to generalize

Machens et al, BMC Infect. Dis.(2013)

Evaluating the representations?



+

Epidemiological model

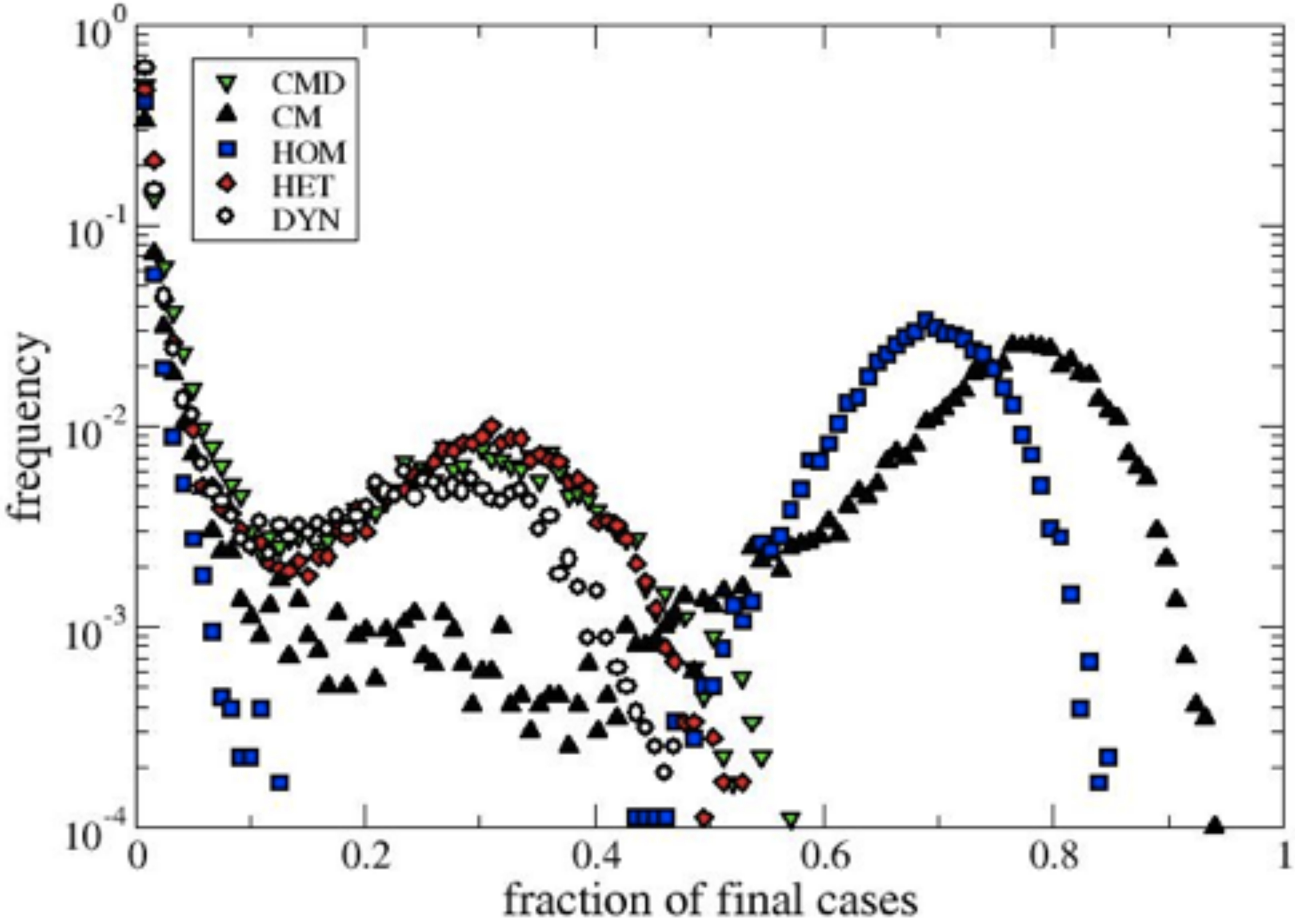
Susceptible → Exposed → Infectious → Recovered

- Evaluation of :
  - Extinction probability
  - Attack rate
  - Role of initial seed
  - Attack rate for each group

- Comparison with most realistic DYN representation

Machens et al, BMC Infect. Dis.(2013)

SEIR simulation results

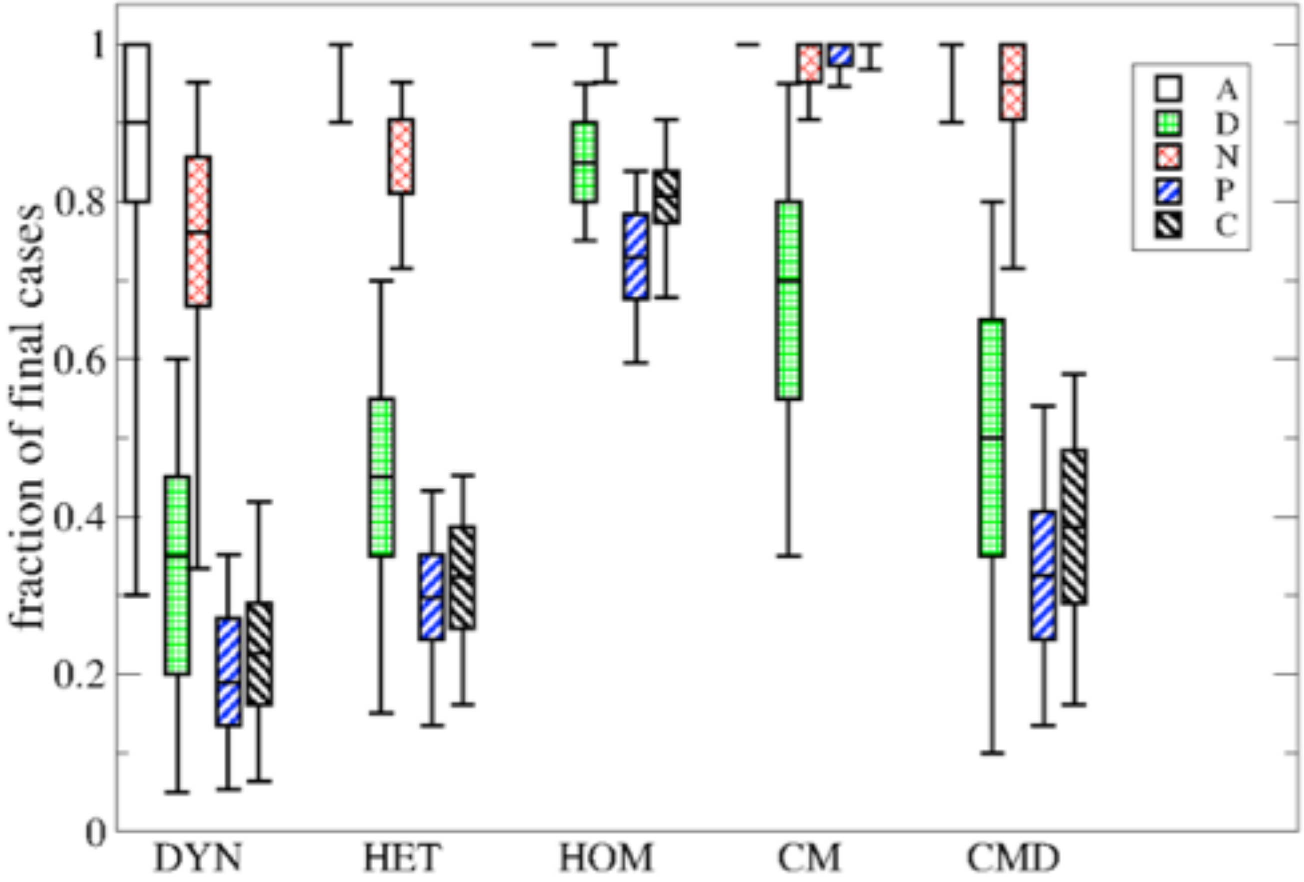


Importance of heterogeneities of contact durations

Machens et al, BMC Infect. Dis.(2013)

SEIR simulation results

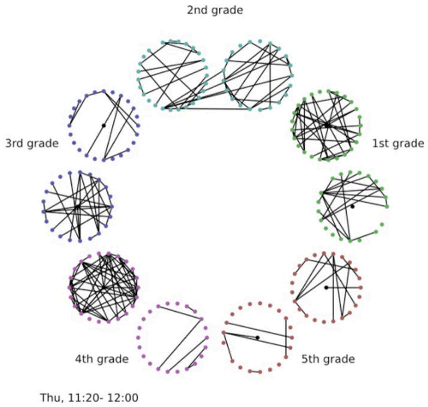
Attack rate by groups (for AR > 10%)



Machens et al, BMC Infect. Dis.(2013)

> Use of data-driven simulations

Scenarii evaluation



+


epidemic model

S ▶ E ▶ I ▶ R

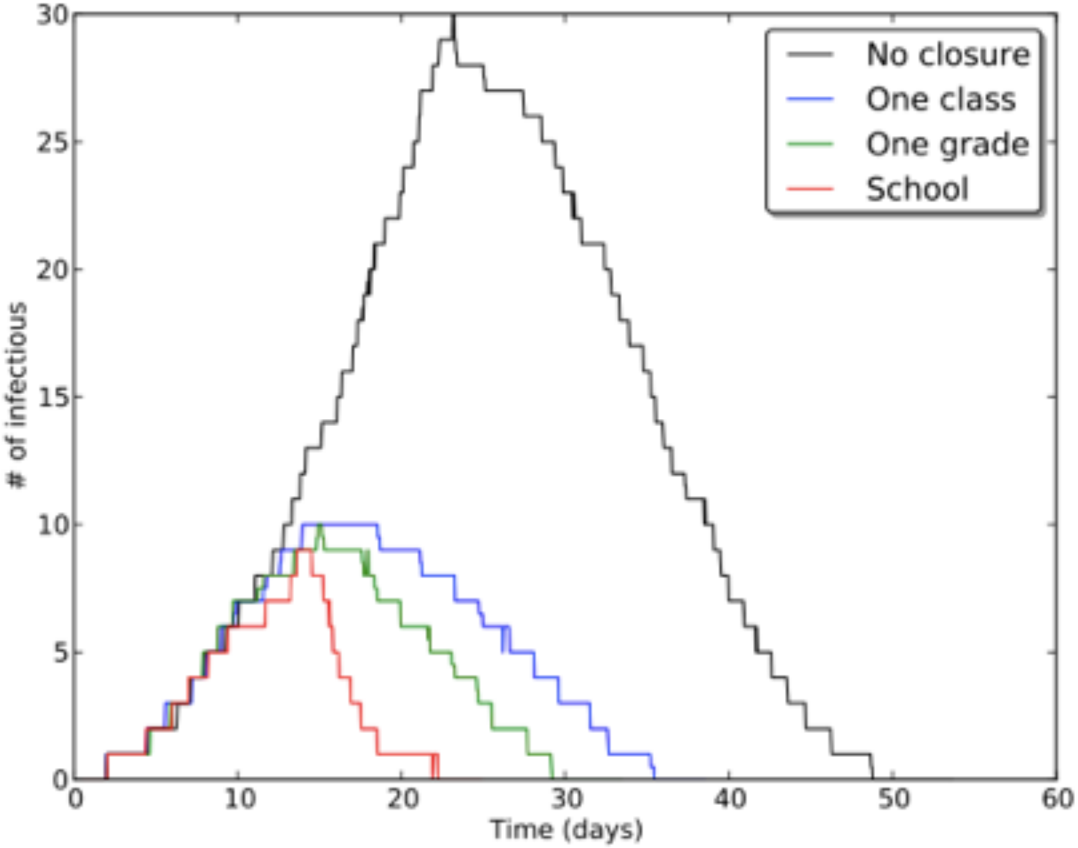
An example: SEIR + school Which containment strategies?

Model:
-SEIR with asymptomatics
-contact data as proxy for possibility of transmission inside school
-when children are out of school: residual homogeneous risk of contamination by contact with population

Containment strategies (suggested by the data):
-detection and subsequent isolation of symptomatic individuals
-whenever symptomatic individuals are detected (more than a given threshold), **closure of**
    (i) class
    (ii) class + most connected other class (same grade)
    (iii) whole school

Gemmetto, Barrat, Cattuto, in preparation

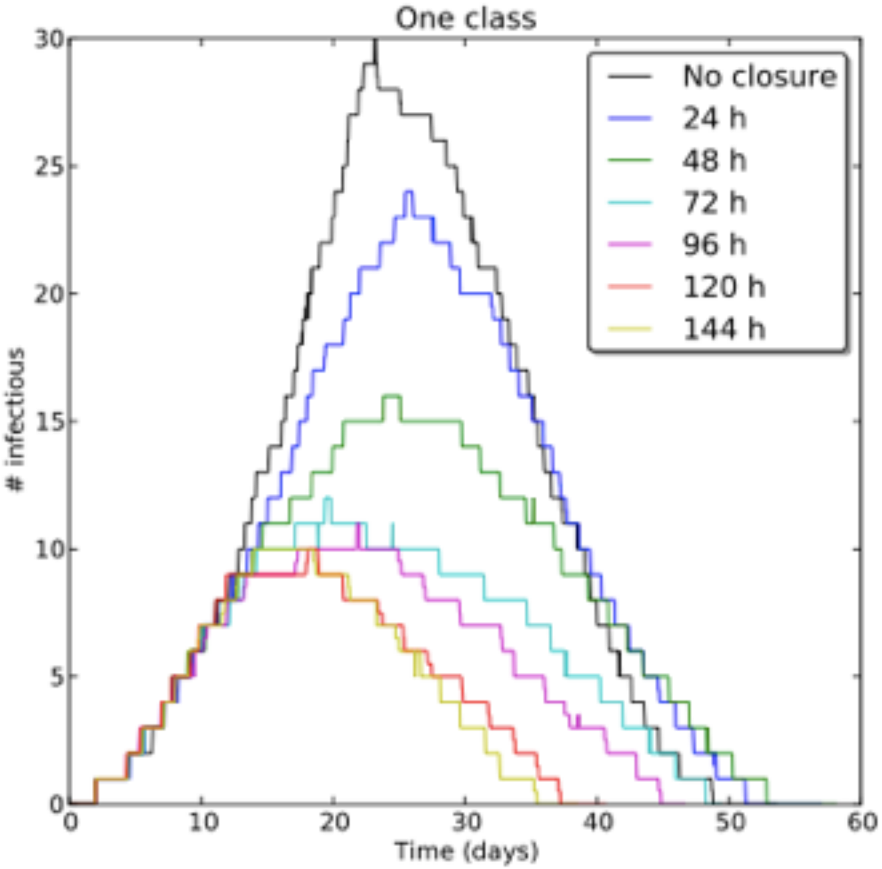
An example: SEIR + school
Which containment strategies?

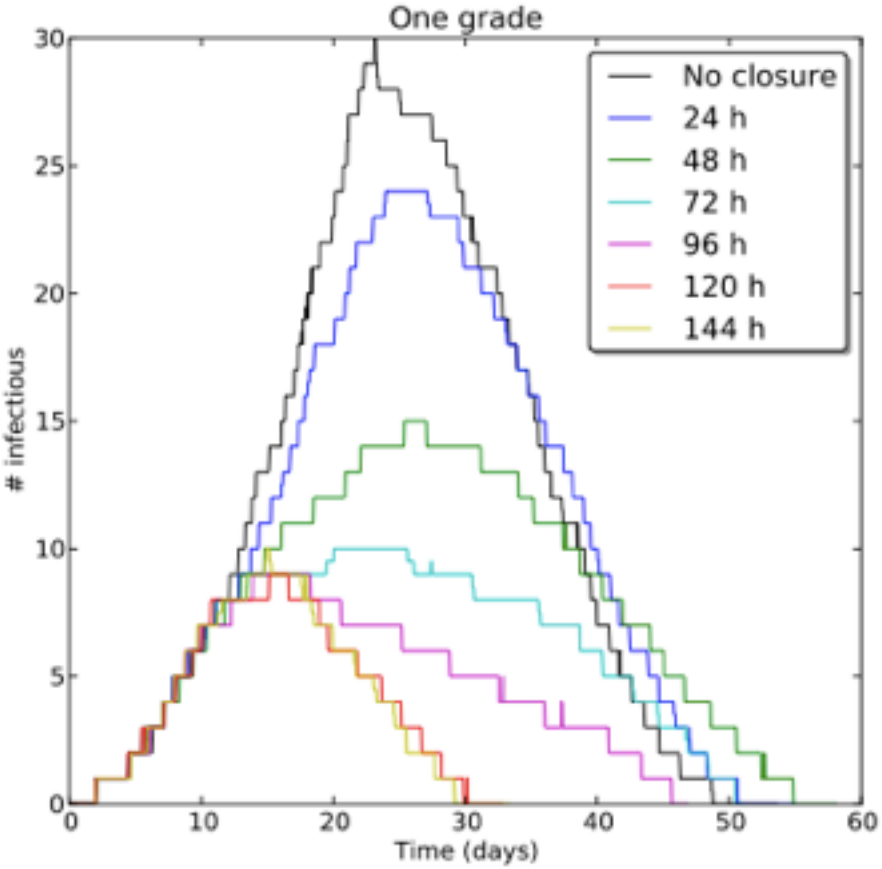


Average over cases
with AR > 10%

Gemmetto, Barrat, Cattuto, in preparation

An example: SEIR + school
Which containment strategies?

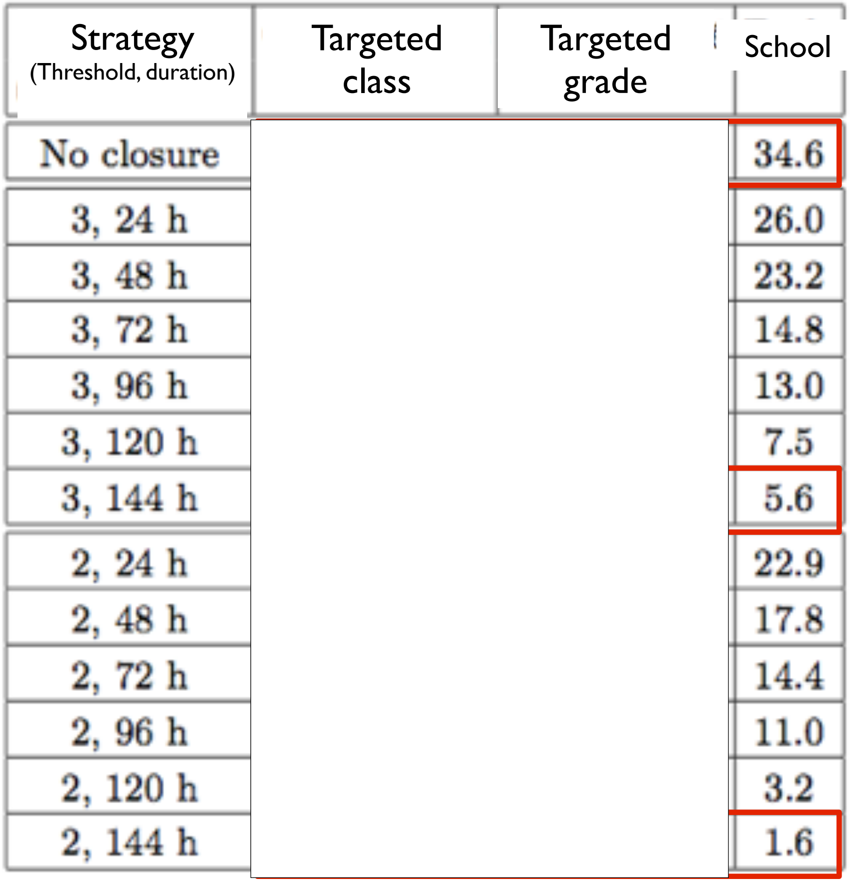




Average over cases
with AR > 10%

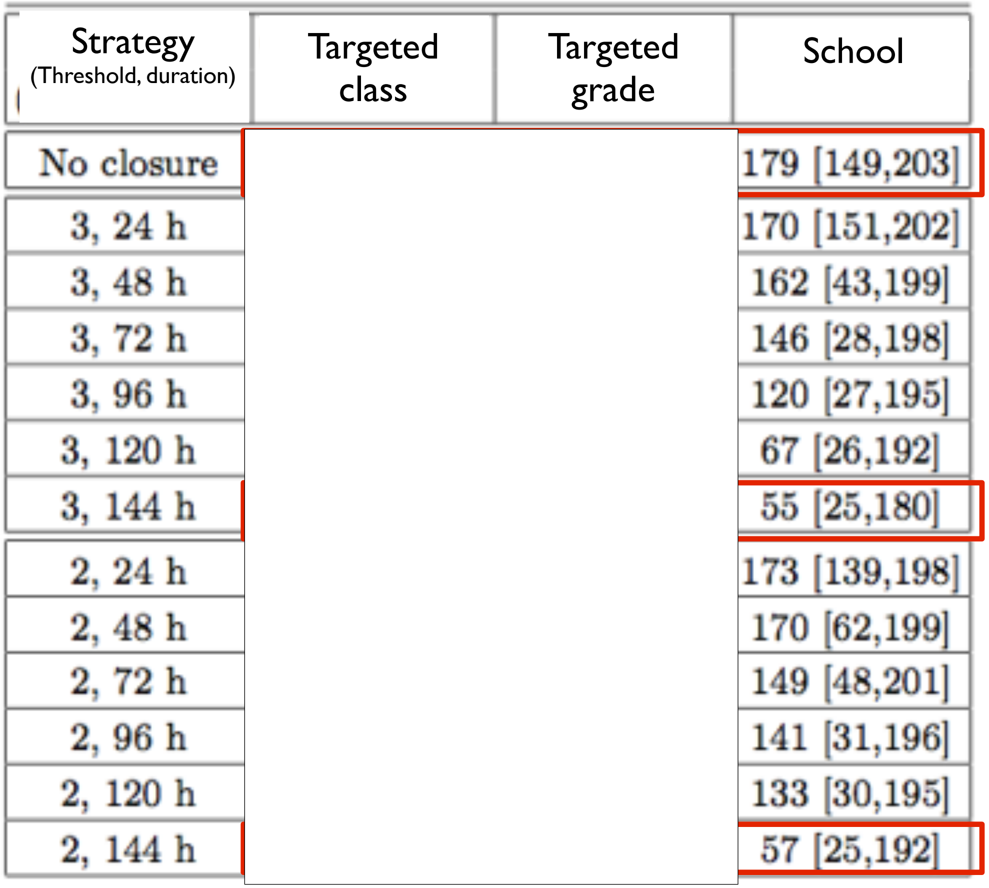
Gemmetto, Barrat, Cattuto, in preparation

Which containment strategies?
Comparing class/grade/school closure



Strategy (Threshold, duration)	Targeted class	Targeted grade	School
No closure			34.6
3, 24 h			26.0
3, 48 h			23.2
3, 72 h			14.8
3, 96 h			13.0
3, 120 h			7.5
3, 144 h			5.6
2, 24 h			22.9
2, 48 h			17.8
2, 72 h			14.4
2, 96 h			11.0
2, 120 h			3.2
2, 144 h			1.6

Probability that AR < 10%

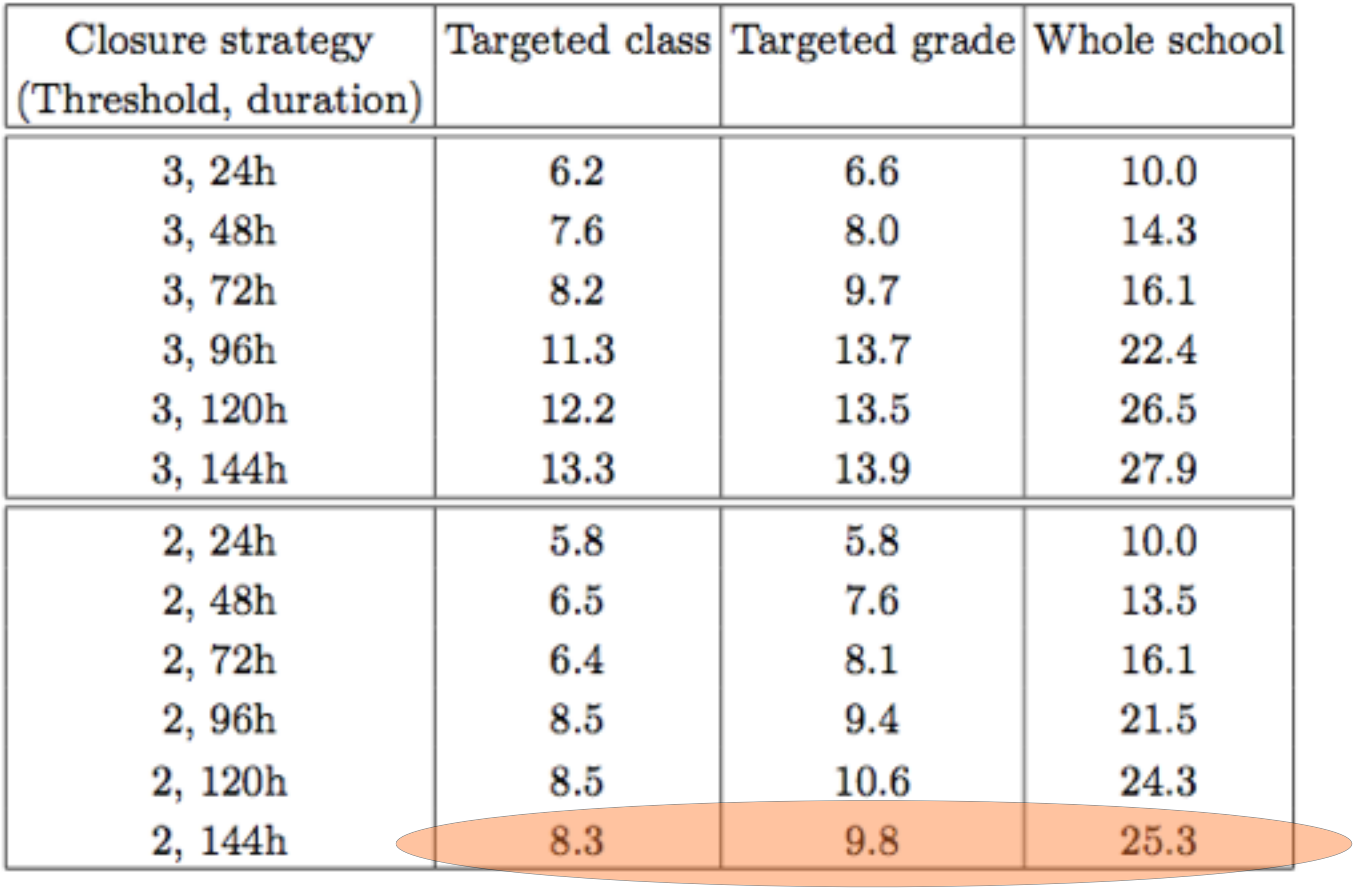


Strategy (Threshold, duration)	Targeted class	Targeted grade	School
No closure			179 [149,203]
3, 24 h			170 [151,202]
3, 48 h			162 [43,199]
3, 72 h			146 [28,198]
3, 96 h			120 [27,195]
3, 120 h			67 [26,192]
3, 144 h			55 [25,180]
2, 24 h			173 [139,198]
2, 48 h			170 [62,199]
2, 72 h			149 [48,201]
2, 96 h			141 [31,196]
2, 120 h			133 [30,195]
2, 144 h			57 [25,192]

Average AR when AR > 10%

Gemmetto, Barrat, Cattuto, in preparation

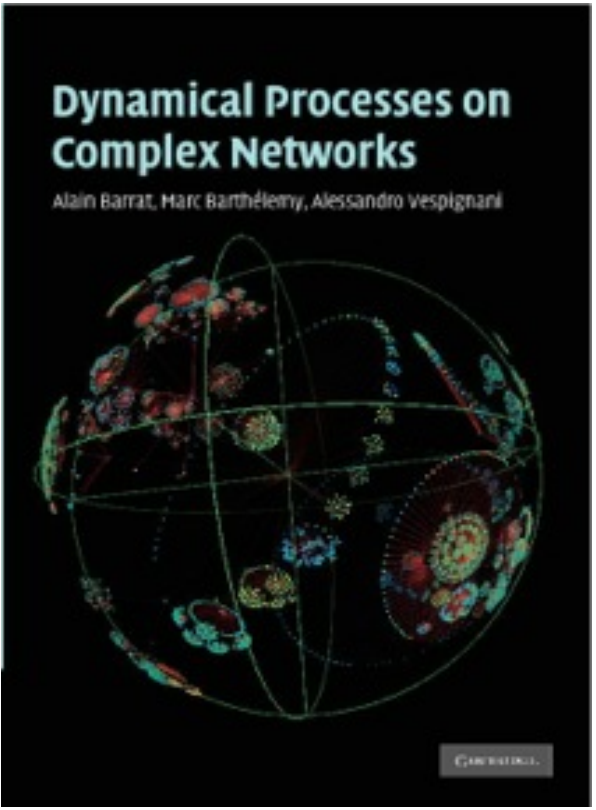
An example: SEIR + school
Which containment strategies?



Closure strategy (Threshold, duration)	Targeted class	Targeted grade	Whole school
3, 24h	6.2	6.6	10.0
3, 48h	7.6	8.0	14.3
3, 72h	8.2	9.7	16.1
3, 96h	11.3	13.7	22.4
3, 120h	12.2	13.5	26.5
3, 144h	13.3	13.9	27.9
2, 24h	5.8	5.8	10.0
2, 48h	6.5	7.6	13.5
2, 72h	6.4	8.1	16.1
2, 96h	8.5	9.4	21.5
2, 120h	8.5	10.6	24.3
2, 144h	8.3	9.8	25.3

Cost in number of lost class-days

Gemmetto, Barrat, Cattuto, in preparation



Dynamical Processes on Complex Networks

Alain Barrat, Marc Barthélemy, Alessandro Vespignani

Physics Reports 519 (2012) 97–125

Contents lists available at SciVerse ScienceDirect

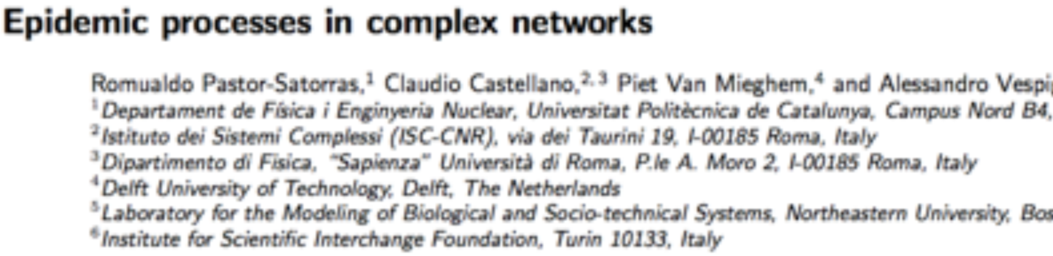
Physics Reports

journal homepage: www.elsevier.com/locate/physrep

Temporal networks

Petter Holme[a,b,c,*], Jari Saramäki[d]

[a] IceLab, Department of Physics, Umeå University, 901 87 Umeå, Sweden
[b] Department of Energy Science, Sungkyunkwan University, Suwon 440–746, Republic of Korea
[c] Department of Sociology, Stockholm University, 106 91 Stockholm, Sweden
[d] Department of Biomedical Engineering and Computational Science, School of Science, Aalto University, 00076 Aalto, Espoo, Finland



Epidemic processes in complex networks

Romualdo Pastor-Satorras,[1] Claudio Castellano,[2,3] Piet Van Mieghem,[4] and Alessandro Vespi
[1] Departament de Física i Enginyeria Nuclear, Universitat Politècnica de Catalunya, Campus Nord B4,
[2] Istituto dei Sistemi Complessi (ISC-CNR), via dei Taurini 19, I-00185 Roma, Italy
[3] Dipartimento di Fisica, "Sapienza" Università di Roma, P.le A. Moro 2, I-00185 Roma, Italy
[4] Delft University of Technology, Delft, The Netherlands
[5] Laboratory for the Modeling of Biological and Socio-technical Systems, Northeastern University, Bos
[6] Institute for Scientific Interchange Foundation, Turin 10133, Italy

SocioPatterns.org